

On the Power-Law Tails of Vote Distributions in Proportional Elections

Filippo Palombi^{a*} and Simona Toti^b

^aENEA – Italian National Agency for New Technologies, Energy and
Sustainable Economic Development, *Via Enrico Fermi 45, 00044 Frascati – Italy*

^bISTAT – Istituto Nazionale di Statistica, *Via Cesare Balbo 16, 00184 Rome – Italy*

January 27, 2016

Abstract

In proportional elections with open lists the excess of preferences received by candidates with respect to the list average is known to follow a universal lognormal distribution. We show that lognormality is broken provided preferences are conditioned to lists with many candidates. In this limit power-law tails emerge. We study the large-list limit in the framework of a quenched approximation of the word-of-mouth model introduced by Fortunato and Castellano (Phys. Rev. Lett. 99(13):138701, 2007), where the activism of the agents is mitigated and the noise of the agent-agent interactions is averaged out. Then we argue that our analysis applies *mutatis mutandis* to the original model as well.

1 Introduction

It is well known that vote distributions in proportional elections display scaling and universality features. The discovery is reported in an inspiring paper by Fortunato and Castellano (FC) [1] and goes as follows. In a country implementing proportional elections with *open* lists, each party presents lists of candidates in one or more electoral districts. Let Q , v , N denote respectively the number of candidates in a given list, the number of preferences assigned to a given candidate in the list and the sum of preferences assigned to all of them. The adjective *open* means that Q is not fixed. The variable $x = vQ/N$ measures the excess of votes assigned to a candidate with respect to the average competitor in the same list. FC show that the probability density function (*p.d.f.*) $\mathcal{F}_{\text{EMP}}(x)$ obtained from the empirical data is independent of Q and N (FC scaling). In addition, they show that $\mathcal{F}_{\text{EMP}}(x)$ is identical in different countries and years and that it is remarkably well fitted by a lognormal distribution $\ln \mathcal{N}(\mu, \sigma^2)$ with $\mu = -0.54$ and $\sigma^2 = 1.08$.

In order to understand the origin of this amazing result, the same authors propose a cascade model based on word of mouth, where agents supporting a given candidate strive to persuade their undecided acquaintances to become supporters of the same candidate¹. The model assumes a network made of Q

*Corresponding author. E-mail: filippo.palombi@enea.it

¹In a more recent paper [2] Burghardt, Rand and Girvan consider a contagion-like model, which is equally able to reproduce the universal lognormal distribution.

independent rooted trees, with candidates sitting on the roots and agents on the vertices. Each vertex has a random number k of *children*, with k following a power-law distribution $\pi(k) \propto k^{-\alpha}$ (it is understood that $\alpha > 2$ and $k \geq k_{\min} \geq 1$). At time $t = 0$ candidates are the only persuaded agents. Each of them interacts with his/her own neighbours and tries to convince them one by one to vote for him/her. A single interaction may have a positive or negative outcome with probabilities r and $1 - r$. At a generic (discrete) time $t > 0$, each persuaded agent in the network tries to convince the undecided neighbours according to the same stochastic rule. The process goes on until the overall number of persuaded agents on all trees equals N . The final number of preferences assigned to each candidate coincides with the final number of persuaded agents on the corresponding tree. Iterating this process (network generation + word-of-mouth spreading) many times and counting preferences in each iteration yields the conditional distribution $F_{\text{FC}}(x|Q, N)$.

Let $X_n^{(i)}$ be the number of vertices on the n th level of the i th tree in a given iteration of the algorithm. Since the network is static, $X_n^{(i)}$ is known at time $t = 0$ for all i and n . Let also $\mathcal{P}_n^{(i)}(t)$ and $\mathcal{U}_n^{(i)}(t)$ denote respectively the subsets of persuaded and undecided agents on the n th level of the i th tree at time t , hence $|\mathcal{P}_n^{(i)}(t)| + |\mathcal{U}_n^{(i)}(t)| = X_n^{(i)}$. The dynamics of the number of persuaded agents $Y_n^{(i)}(t) = |\mathcal{P}_n^{(i)}(t)|$ is quantified by the equations

$$\begin{cases} Y_0^{(i)}(t) = 1 & \text{for all } t \geq 0, \\ Y_n^{(i)}(t) = 0 & \text{if } t < n, \\ Y_n^{(i)}(t) = Y_n^{(i)}(t-1) + \sum_{a \in \mathcal{U}_n^{(i)}(t-1)} \xi_a \mathbf{1}_{\mathcal{P}_{n-1}^{(i)}(t-1)}(a') & (a' \text{ is the parent vertex of } a), \text{ if } t \geq n, \end{cases} \quad (1.1)$$

where $\{\xi_a\}$ are i.i.d. Bernoulli variables with success probability r and $\mathbf{1}_A(x)$ represents the indicator function of A , *i.e.* $\mathbf{1}_A(x) = 1$ if $x \in A$ and $\mathbf{1}_A(x) = 0$ otherwise. In this framework, the number of preferences assigned to the i th candidate at time t is given by

$$V_{\text{FC}}^{(i)}(t) = \sum_{s=0}^t Y_s^{(i)}(t). \quad (1.2)$$

Eq. (1.1) shows at a glance that the FC algorithm is a Markov process. It also shows that the algorithm is not an ordinary branching process, as FC firstly notice in their paper: once an agent gets persuaded, he/she keeps trying to convince his/her undecided acquaintances at all subsequent times. Indeed, the inequality $Y_n^{(i)}(t+1) \geq Y_n^{(i)}(t)$ holds in general for all n and t .

The conditional distribution $F_{\text{FC}}(x|Q, N)$ cannot be directly compared with $\mathcal{F}_{\text{EMP}}(x)$. Since $Q/N \leq x \leq Q$, it follows that $F_{\text{FC}}(x|Q, N) = 0$ for $x < Q/N$ and $x > Q$. Hence, $F_{\text{FC}}(x|Q, N)$ depends explicitly upon Q and N . As such it violates the FC scaling. To get rid of this, FC run their algorithm for all (Q, N) occurring in the empirical datasets and convolve the resulting curves. More precisely, let $p(Q, N)$ denote the empirical probability of (Q, N) , namely

$$p(Q, N) = \frac{\text{no. of pairs } (Q, N) \text{ in the empirical datasets}}{\text{no. of all pairs } (Q', N') \text{ in the empirical datasets}}. \quad (1.3)$$

The FC scaling distribution is then given by

$$\mathcal{F}_{\text{FC}}(x) = \sum_{Q, N} p(Q, N) F_{\text{FC}}(x|Q, N). \quad (1.4)$$

The agreement of $\mathcal{F}_{\text{FC}}(x)$ with $\mathcal{F}_{\text{EMP}}(x)$ is just perfect provided the model parameters are set to $(\alpha, r, k_{\min}) = (2.45, 0.25, 10)$. Convolution of $F_{\text{FC}}(x|Q, N)$ with $p(Q, N)$ is truly essential for reproducing the empirical distribution: $F_{\text{FC}}(x|Q, N)$ alone is known to deliver a power-law right tail as $Q \rightarrow \infty$; it is only its weighted

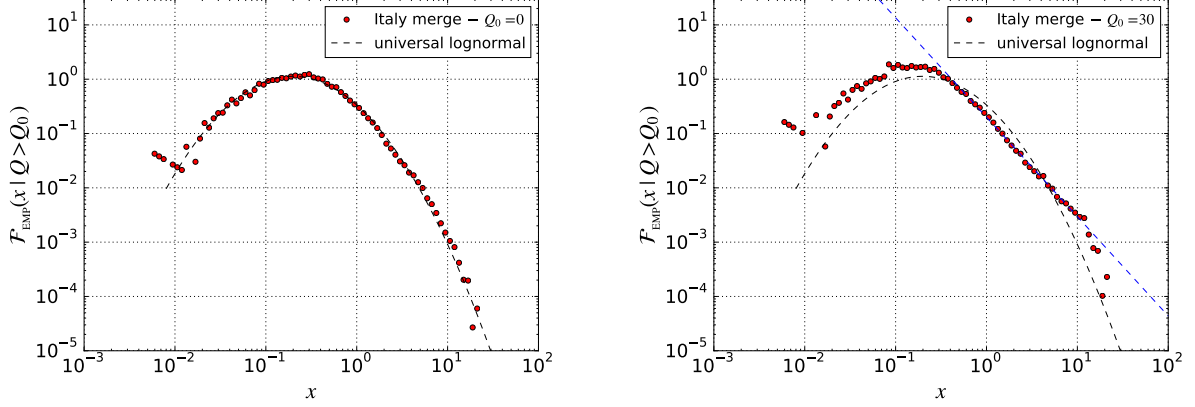


Figure 1 – (Left) complete empirical distribution of the excess-of-votes variable x obtained by merging data of the Italian elections held in 1958, 1972, 1976, 1979, 1987; (Right) conditional empirical distribution of x given $Q > Q_0 = 30$ obtained by merging the same datasets.

average over (Q, N) that yields the observed lognormal behaviour. The presence of power-law tails at fixed (Q, N) has been regarded so far as a problem of the model [3].

On a closer inspection, power-law structures emerge from the empirical data when focusing on a specific subset of them. In place of $\mathcal{F}_{\text{EMP}}(x)$ we introduce the conditional distribution $\mathcal{F}_{\text{EMP}}(x|Q > Q_0)$, with Q_0 properly chosen. This is obtained by selecting data which belong only to lists with $Q > Q_0$. An example is given in Fig. 1, where the empirical distribution of x , obtained by merging data from the Italian elections held from 1958 to 1987, is shown for $Q_0 = 0$ (complete distribution) and $Q_0 = 30$. Lognormality is manifestly broken in the latter distribution, which is instead characterized by a power-law right tail with final collapse at $x \lesssim Q_0$. With reference to the group U of countries studied in ref. [4], it must be added that: *i*) a similar behaviour is displayed by data of Polish and Estonian elections; *ii*) Danish data have no lists with $Q > Q_0 = 30$; *iii*) Finnish data still display a lognormal conditional distribution given $Q_0 = 30$, hence they represent a puzzling exception; *iv*) Italian data are those with the largest statistics for $Q > Q_0 = 30$; this is why we plot just them in Fig. 1.

We propose an interpretation of this effect. We first notice that by construction $\mathcal{F}_{\text{EMP}}(x|Q > Q_0)$ fulfills the inequality

$$\mathcal{F}_{\text{EMP}}(x) < \mathcal{F}_{\text{EMP}}(x|Q > Q_0), \quad \text{for } x > Q_0 \quad (1.5)$$

(we refer the reader to app. A for a simple proof of eq. (1.5)). As a consequence, the opposite inequality $\mathcal{F}_{\text{EMP}}(x) > \mathcal{F}_{\text{EMP}}(x|Q > Q_0)$ must hold somewhere in the complementary region $x \leq Q_0$, since both $\mathcal{F}_{\text{EMP}}(x)$ and $\mathcal{F}_{\text{EMP}}(x|Q > Q_0)$ are correctly normalized as *p.d.f.*'s. Secondly, $\mathcal{F}_{\text{EMP}}(x|Q > Q_0)$ also can be represented as a convolution of conditional distributions $\mathcal{F}_{\text{EMP}}(x|Q, N)$ with $Q > Q_0$, *viz.*

$$\mathcal{F}_{\text{EMP}}(x|Q > Q_0) = \sum_{Q > Q_0} \sum_N p(Q, N|Q > Q_0) \mathcal{F}_{\text{EMP}}(x|Q, N), \quad (1.6)$$

where the conditional probability $p(Q, N|Q > Q_0)$ is defined for $Q > Q_0$ by

$$p(Q, N|Q > Q_0) = \frac{\text{no. of pairs } (Q, N) \text{ in the empirical datasets}}{\text{no. of all pairs } (Q', N') \text{ with } Q' > Q_0 \text{ in the empirical datasets}}. \quad (1.7)$$

Although we have insufficient statistics for measuring $\mathcal{F}_{\text{EMP}}(x|Q, N)$, we know that it vanishes for $x > Q$ by construction. Hence, we expect it to be contaminated by truncation artefacts for $x \lesssim Q$, amounting to a faster decay just to the left of the truncation point, than we would observe for larger values of Q .

Independently of how these local distortions are compensated along the domain of the distribution², their overall impact is modest provided Q is sufficiently large. Therefore, it is reasonable to assume $F_{\text{EMP}}(x|Q, N)$ to become independent of Q for $x \ll Q$ as $Q \rightarrow \infty$. If this is correct, it follows that $\mathcal{F}_{\text{EMP}}(x|Q > Q_0)$ is almost insensitive to its defining convolution, *i.e.* for $x \ll Q_0$ we must have

$$\begin{aligned} \mathcal{F}_{\text{EMP}}(x|Q > Q_0) &= \sum_{Q > Q_0} \sum_N p(Q, N|Q > Q_0) F_{\text{EMP}}(x|Q, N) \\ &\simeq \sum_N F_{\text{EMP}}(x|Q_0, N) \sum_{Q > Q_0} p(Q, N|Q > Q_0) \\ &= \sum_N F_{\text{EMP}}(x|Q_0, N) p(N|Q > Q_0) \simeq F_{\text{EMP}}(x|Q_0, \bar{N}), \end{aligned} \quad (1.8)$$

with $\bar{N} = \mathbb{E}[N|Q > Q_0] \gg 1$ (there is always a positive correlation between Q and N in the empirical data, because parties present larger lists of candidates in larger constituencies). If we believe in the FC model, then $F_{\text{EMP}}(x|Q_0, \bar{N})$ is well reproduced by $F_{\text{FC}}(x|Q_0, \bar{N})$. Accordingly, $F_{\text{EMP}}(x|Q_0, \bar{N})$ has a power-law tail for $1 \lesssim x \lesssim Q_0$ and we can conclude that $\mathcal{F}_{\text{EMP}}(x|Q > Q_0)$ also has a power-law tail in the same region.

The above considerations suggest to regard the power-law tail of $F_{\text{FC}}(x|Q, N)$ not as an unphysical attribute of the model, but – quite the opposite – as an essential feature of the phenomenon it describes. Since the complete tail of the distribution emerges only as $Q \rightarrow \infty$, we are naturally led to consider the large-list limit

$$\mathcal{F}_{\text{FC}}^*(x) = \lim_{Q \rightarrow \infty} \lim_{N \rightarrow \infty} F_{\text{FC}}(x|Q, N). \quad (1.9)$$

Quite interestingly, this limit represents another realization of the FC scaling. Unlike eq. (1.4), it does not require any input from the empirical data. The ultimate reason why there exist two different ways to build a scaling function out of $F_{\text{FC}}(x|Q, N)$ is not clear.

Aim of the present paper is to investigate $\mathcal{F}_{\text{FC}}^*(x)$ and its corrections at finite Q . Despite the significance of the FC model, the analytic structure of $F_{\text{FC}}(x|Q, N)$ has never been studied in the literature so far, to the best of our knowledge. The analysis we present here is the first attempt to go beyond numerical simulations. Given the complexity of the FC model, in sect. 2 we consider a quenched approximation of it, where the word of mouth generates Galton–Watson trees with Mandelbrot offspring distribution. Albeit simpler, the quenched model is still in good agreement with the empirical data. Based on the exponential scaling of Galton–Watson trees, in sect. 3 we work out the vote distribution in the large-list limit and study how it changes under variations of the stopping rule of the algorithm. In sect. 4 we go back to the original model and show that an analogous solution applies to it *mutatis mutandis*. Specifically, we find that, similar to Galton–Watson trees, also trees resulting from eq. (1.1) scale exponentially as $t \rightarrow \infty$ and we calculate their growth rate. This is sufficient for studying the FC model in the large-list limit. Finally, in sect. 5 we derive an asymptotic estimate of the power-law tail of the vote distribution in the quenched model. This applies with rather good approximation to the FC model too. We draw our conclusions in sect. 6.

2 Quenched word-of-mouth model

As mentioned above, the major complication of the FC model is represented by the activism of the agents, who repeatedly endeavor to persuade their undecided acquaintances. A great simplification is achieved if we prevent them from insisting after their first attempt. This turns the model into a standard Galton–Watson

²They might be absorbed locally in a region close to the boundary $x \lesssim Q$, locally in a different region within the domain of the distribution or globally through a dilution along the whole domain.

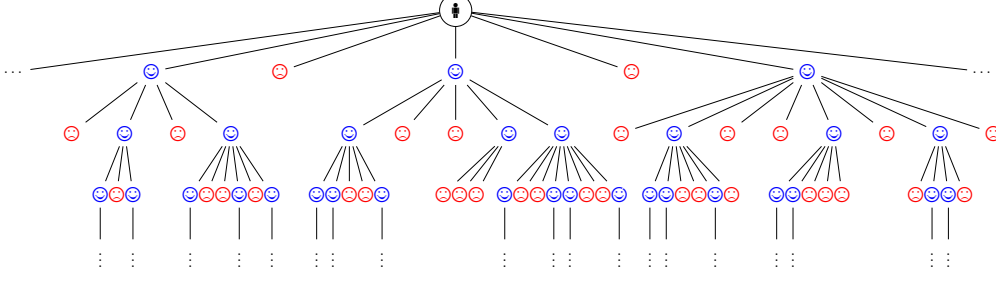


Figure 2 – A sample tree with agents freezing after one Bernoulli interaction with their children. Voters (smiling agents) branch according to $p_B(\ell)$. Non-voters (frowning agents) are not regarded as vertices: they are included in the picture only for illustrative purposes.

process, where the time variable t comes to be identified with the level variable n . In this approximation candidates interact with their neighbours at time $t = n = 0$, a single interaction being represented by a Bernoulli trial with success probability r . Hence they freeze, while persuaded agents on level $t = n = 1$ interact in turn with their acquaintances lying on the next level, according to the same stochastic prescription. Then they freeze too and so on and so forth. If $Z_t^{(i)}$ denotes the number of persuaded agents on the t th level of the i th tree³, the word-of-mouth spreading is described by the simpler equations

$$\begin{cases} Z_0^{(i)} = 1, \\ Z_{t+1}^{(i)} = \sum_{j=0}^{Z_t^{(i)}} \xi_{t,j}^{(i)}, \quad t \geq 0, \\ \{\xi_{t,j}^{(i)}\} \text{ are i.i.d. variables with } \xi_{t,j}^{(i)} \sim p_B(\ell) = C^{-1} \sum_{k=\max\{\ell, k_{\min}\}}^{\infty} \binom{k}{\ell} \frac{r^\ell (1-r)^{k-\ell}}{k^\alpha}, \end{cases} \quad (2.1)$$

where $\ell \geq 0$ and $C = \sum_{\ell=0}^{\infty} \sum_{k=\max\{\ell, k_{\min}\}}^{\infty} \binom{k}{\ell} \frac{r^\ell (1-r)^{k-\ell}}{k^\alpha}$ is the normalization constant of the offspring distribution $p_B(\ell)$. The subindex B (roman font) reminds us that the agent-agent interaction is a Bernoulli trial. A sample tree generated according to the above procedure is depicted in Fig. 2. We observe that $p_B(\ell)$ is precisely the probability of contacting k agents with probability $\pi(k) \propto k^{-\alpha}$ under the condition that $k \geq k_{\min}$ and of convincing a fraction $\ell \leq k$ of them, each with probability r , in all possible ways. For later convenience, we let $m_B = \sum_{\ell=0}^{\infty} \ell p_B(\ell)$ denote the average offspring of a vertex. Moreover, in this framework the number of preferences assigned to the i th candidate up to level t is given by

$$V_t^{(i)} = \sum_{s=0}^t Z_s^{(i)}. \quad (2.2)$$

The statistical properties of $Z_t^{(i)}$ (in particular the tail of its distribution) are completely determined by $p_B(\ell)$. Unfortunately, despite the simplification operated on the original process, the structure of $p_B(\ell)$ is still too complex to allow for analytic calculations. Hence, we proceed to a further reduction of complexity. We notice that if a persuaded agent has many contacts k and interacts with them separately via Bernoulli trials with success probability r , on average he/she persuades a fraction rk of them. Indeed, the relation

$$r \sum_{k=\ell}^{\infty} \binom{k}{\ell} \frac{r^\ell (1-r)^{k-\ell}}{k^\alpha} \stackrel{\ell \rightarrow \infty}{\approx} \frac{r^\alpha}{\ell^\alpha} \quad (2.3)$$

holds asymptotically. Inspired by this equation, we introduce the integer variable $n_{\min} = \max\{1, \lfloor rk_{\min} \rfloor\}$

³When not needed, we shall drop the upper index of $Z_t^{(i)}$ and simply write Z_t in place of it.

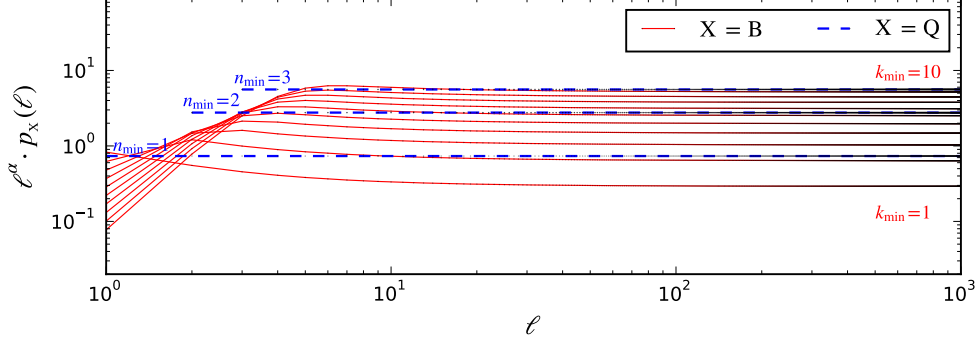


Figure 3 – Probabilities $p_B(\ell)$ (continuous lines) and $p_Q(\ell)$ (dashed lines) for $(\alpha, r) = (2.45, 0.25)$.

and in place of eq. (2.1) we consider the Galton–Watson process

$$\begin{cases} Z_0^{(i)} = 1, \\ Z_{t+1}^{(i)} = \sum_{j=0}^{Z_t^{(i)}} \xi_{t,j}^{(i)}, \quad t \geq 0, \\ \{\xi_{t,j}^{(i)}\} \text{ are i.i.d. variables with } \xi_{t,j}^{(i)} \sim p_Q(n) = \frac{1}{\zeta(\alpha, n_{\min})} \frac{1}{n^\alpha}, \end{cases} \quad (2.4)$$

where $n \geq n_{\min}$ and $\zeta(\alpha, n_{\min}) = \sum_{n=0}^{\infty} \frac{1}{(n+n_{\min})^\alpha}$ denotes the Hurwitz ζ -function. The offspring distribution $p_Q(n)$ is known in the literature as the *shifted power law* or Mandelbrot distribution [5]. In the sequel we focus on eq. (2.4) and refer to it as the *quenched* model, with the understanding that it has been obtained by quenching both the activism of the agents and the fluctuations of the single interactions characterizing the original FC model. The subindex Q (roman font) in $p_Q(\ell)$ stands for *quenched*.

In Fig. 3 we compare the offspring distributions $p_B(\ell)$ and $p_Q(n)$ for $\alpha = 2.45$, $r = 0.25$ and several choices of k_{\min} and n_{\min} just to stress that: *i*) $p_Q(n)$ undergoes discrete jumps as $n_{\min} \rightarrow n_{\min} + 1$; *ii*) apart from the overall normalization, the behaviour of the two distributions essentially differs only for $\ell \sim O(1)$.

We let \mathcal{T} denote the space of trees (elementary events) and \mathcal{M} the additive class of subsets of \mathcal{T} (events). The measure theory in \mathcal{T} has been first discussed in ref. [6], to which we refer the reader for details. Eqs. (2.1) and (2.4) induce two different probability spaces on \mathcal{T} , namely

- eq. (2.1) $\rightarrow \mathcal{P}_B(\alpha, r, k_{\min}) = (\mathcal{T}, \mathcal{M}, \mathbb{P}_B\{\cdot | \alpha, r, k_{\min}\})$;
- eq. (2.4) $\rightarrow \mathcal{P}_Q(\alpha, n_{\min}) = (\mathcal{T}, \mathcal{M}, \mathbb{P}_Q\{\cdot | \alpha, n_{\min}\})$.

Eq. (1.1) induces a family of probability spaces $\mathcal{P}_{FC,t}(\alpha, r, k_{\min}) = (\mathcal{T}_t, \mathcal{M}_t, \mathbb{P}_{FC,t}\{\cdot | \alpha, r, k_{\min}\})$ parametrized by $t = 0, 1, 2, \dots$, where \mathcal{T}_t is the space of finite trees with at most t levels and \mathcal{M}_t is the additive class of subsets of \mathcal{T}_t . For simplicity, in the sequel we use the notation $\mathbb{P}\{R\}$ with no subindices and/or parameters to denote the probability of $R \in \mathcal{M}$ whenever the context makes it clear to which specific measure we refer.

Remark 2.1. The network of acquaintances in the FC model is a collection of disjoint and i.i.d. trees, each belonging to $\mathcal{P}_Q(\alpha, k_{\min})$, *i.e.* $X_{n+1}^{(i)} = \sum_{j=0}^{X_n^{(i)}} \xi_{n,j}^{(i)}$ with $\xi_{n,j}^{(i)} \sim p_Q(n)|_{n_{\min} \rightarrow k_{\min}}$. ■

So far we have discussed the structure of single trees. In the quenched word-of-mouth model we generate trees in parallel level by level: we first draw a realization of $(Z_1^{(i)})_{i=1}^Q$, hence we draw a realization of $(Z_2^{(i)})_{i=1}^Q$, *etc.* As $t \rightarrow \infty$, the algorithm generates Q -tuples of trees, *i.e.* elements of the Cartesian product

$$\mathfrak{T} = \underbrace{\mathcal{T} \times \dots \times \mathcal{T}}_{Q \text{ times}}. \quad (2.5)$$

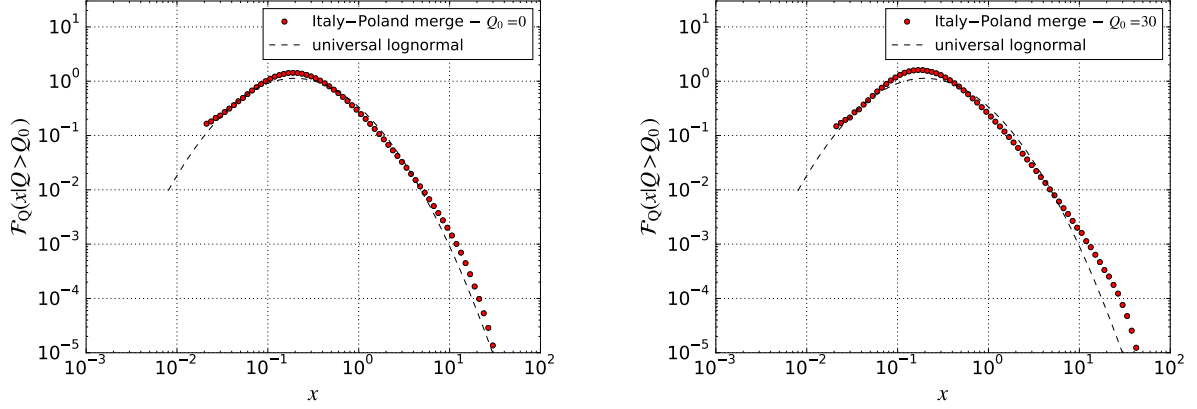


Figure 4 – (Left) complete distribution of excess-of-votes variable x obtained by convolving $F_Q(x|Q, N)$ with the empirical probability $p(Q, N)$; (Right) conditional distribution of x obtained by convolving $F_Q(x|Q, N)$ with the empirical probability $p(Q, N|Q > Q_0)$ and $Q_0 = 30$. In both plots, the empirical probability has been obtained by merging datasets of the Italian elections held in 1958, 1972, 1976, 1979, 1987 and the Polish elections held in 2001, 2005, 2007, 2011.

The set \mathfrak{T} is the space of elementary events of the model. The probability measure on \mathfrak{T} is the product measure obtained in the usual way for independent variables, see ref. [7, ch. IV.6]. In the following, we let $T = (T^{(1)}, \dots, T^{(Q)})$ denote a generic element of \mathfrak{T} . We call T a Q -forest. Incidentally, we observe that the stochastic variable $V_t^{(i)}$ is a function of $T^{(i)}$ alone. Therefore, we write $V_t^{(i)} = V_t(T^{(i)})$ whenever appropriate.

To complete the algorithmic prescription we need to specify a stopping rule to arrest the generation of new levels. In sect. 3 we study, among other things, how the vote distribution depends on this part of the procedure. We consider three possible stopping rules:

- SR1: the process stops as soon as the overall number of vertices generated on all trees equals N . This is the stopping rule adopted by FC in ref. [1]. It has the drawback that precisely when the process stops, some trees have developed down to a certain level t_{exit} , while the others only down to level $t_{\text{exit}} - 1$. In most cases there is a tree $T^{(i)}$ for which $Z_{t_{\text{exit}}}^{(i)}$ is only partially realized. This complicates the analytic description, as we shall see;
- SR2: the process stops as soon as the generation of the earliest level t_{exit} , at which the overall number of vertices on all trees reaches N , is complete for all trees. Hence, all variables $(Z_{t_{\text{exit}}}^{(i)})_{i=1}^Q$ are fully realized. Unfortunately, this rule has the drawback that precisely when the process stops, the overall number M of vertices generated on all trees fluctuates along different iterations of the algorithm, with $M \geq N$;
- SR3: the process stops as soon as the generation of a predefined level t_{exit} is complete for all trees. Similar to the previous case, all variables $(Z_{t_{\text{exit}}}^{(i)})_{i=1}^Q$ are fully realized. The overall number M of vertices generated on all trees fluctuates along different iterations, however M is totally unconstrained in this case.

Iterating the above algorithm many times and counting preferences yields the conditional distribution $F_Q(x|Q, N)$ (under SR1 and SR2) or $F_Q(x|Q, t_{\text{exit}})$ (under SR3). We incidentally notice that under SR1 it is no matter whether trees are randomly or sequentially ordered when the level t_{exit} is generated, provided $F_Q(x|Q, N)$ is averaged over all candidates.

In Fig. 4 we show the distribution $\mathcal{F}_Q(x)$ obtained by convolving the conditional distribution $F_Q(x|Q, N)$ (under SR2) with the empirical probability $p(Q, N)$ obtained by merging the datasets of the Italian elections

$m_Q = \frac{\zeta(\alpha-1, n_{\min})}{\zeta(\alpha, n_{\min})}$		α				
		2.35	2.45	2.55	2.65	2.75
n_{\min}	2	6.04419	5.06405	4.44184	4.01235	3.69847
	3	9.80849	8.20382	7.18374	6.47842	5.96197
	4	13.62090	11.38611	9.96479	8.98148	8.26097

Table 1 – Values of m_Q for various choices of (α, n_{\min}) .

held from 1958 to 1987 jointly with those of the Polish elections held from 2001 to 2011. The curves correspond to a choice of the model parameters $(\alpha, n_{\min}) = (2.45, 3)$. The agreement of the complete distribution (left) with the universal lognormal is not perfect. Yet, it is sufficiently good to let us regard the quenched model as a valid simplification for studying the large-list limit and its corrections at finite Q . The emergence of a power-law right tail for $Q_0 = 30$ (right) is less pronounced than in the empirical data, yet it is still visible.

Remark 2.2. At this point the reader may wonder why the quenched model matches the empirical data so well, in spite of the razor cuts operated on the FC model. To answer this, we notice that from an algorithmic point of view a typical tree $T^{(i)} \in \mathcal{P}_{\text{FC}, t_{\text{exit}}}(\alpha, r, k_{\min})$ looks like a hybrid. If t_{exit} is not too small, persuaded voters $Y_n^{(i)}(t_{\text{exit}})$ on the first levels ($n \ll t_{\text{exit}}$) are close to their maximum possible value $X_n^{(i)}$, since undecided agents have been contacted several times. This part of the tree looks much like typical elements of $\mathcal{P}_Q(\alpha, k_{\min})$. Of note, the latter are much fatter than typical elements of $\mathcal{P}_Q(\alpha, n_{\min})$. By contrast, undecided agents on the last levels ($n \simeq t_{\text{exit}}$) have been contacted just a few times: specifically, those on the t_{exit} th level have undergone just one interaction with their parents. Therefore, this part of the tree resembles more typical elements of $\mathcal{P}_B(\alpha, r, k_{\min})$ and these are similar to typical elements of $\mathcal{P}_Q(\alpha, n_{\min})$. Moreover, the last levels are those which contribute most to the overall number of preferences. We conclude that the quenched model works well because trees in $\mathcal{P}_Q(\alpha, n_{\min})$ are similar to trees in $\mathcal{P}_{\text{FC}, t_{\text{exit}}}(\alpha, r, k_{\min})$ precisely where they should be, *i.e.* on the last levels. Notice that the optimal choice of the FC model parameters yields $rk_{\min} = 2.5$, whereas in the quenched model we have $n_{\min} = 3$. This yields an average compensation of all topological differences. Of course, the above intuitive argument should be substantiated by a quantitative analysis of the growth rate of $Y_n^{(i)}(t)$. We shall discuss this in sect. 4. ■

Let us now examine more closely trees in $\mathcal{P}_Q(\alpha, n_{\min})$. The standard toolbox to study them is the theory of branching processes, discussed for instance in the classical monographs by Harris [8] and Athreya and Ney [9]. The average offspring of a vertex is given by

$$m_Q \equiv \sum_{n=n_{\min}}^{\infty} n p_Q(n) = \frac{\zeta(\alpha-1, n_{\min})}{\zeta(\alpha, n_{\min})} > 1. \quad (2.6)$$

For the reader's convenience, in Table 1 we collect values of m_Q for a handful of plausible choices of (α, n_{\min}) . Owing to $m_Q > 1$, quenched trees are supercritical, *i.e.* on average they fulfill $Z_t \rightarrow \infty$ as $t \rightarrow \infty$. We recall indeed that

$$\mathbb{E}[Z_t] = \mathbb{E}\left[\mathbb{E}\left[\sum_{j=0}^{Z_{t-1}} \xi_{t,j} \mid Z_{t-1}\right]\right] = \mathbb{E}[m_Q Z_{t-1}] = m_Q \mathbb{E}[Z_{t-1}]. \quad (2.7)$$

Since $\mathbb{E}[Z_0] = 1$, it follows that $\mathbb{E}[Z_t] = m_Q^t$. Hence, we see that m_Q is the exponential growth rate of Z_t . We also notice that $m_Q \sim \mathcal{O}(10)$ if α is not too large and n_{\min} not too small. As a consequence $Z_t, Z_{t+1}, Z_{t+2}, \dots$ belong to different orders of magnitude on average. We finally observe that since $p_Q(n) = 0$ for

$n < n_{\min}$, it follows that $\mathbb{P}\{Z_t = 0\} = 0$ for all t^4 . In cases like this, it is appropriate to introduce the variable $W_t = Z_t/m_Q^t$. Obviously, W_t fulfills $\mathbb{E}[W_t] = 1$. Moreover, an argument analogous to eq. (2.7) proves that W_t is a martingale, that is to say it fulfills

$$\mathbb{E}[W_t | W_{t-1}] = W_{t-1}. \quad (2.8)$$

The latter two properties ensure that the sequence $(W_t)_{t \geq 0}$ converges to a finite limit W . In Fig. 5 (top) we show the *p.d.f.* $\phi(w)$ of W obtained from numerical simulations of W_7 performed at $\alpha = 2.45, 2.65, 2.85$ and $n_{\min} = 2, 3, 4$ (simulations suggest that W_t is close to convergence for $t \gtrsim 4$). It is known that the probability generating function of W , namely $G_w(z) = \mathbb{E}[z^W]$, obeys the functional equation

$$G_w(z) = G_Q \left(G_w(z^{1/m_Q}) \right), \quad (2.9)$$

where $G_Q(z) = \mathbb{E}[z^{Z_1}] = \sum_{n=n_{\min}}^{\infty} p_Q(n) z^n$ is the probability generating function of Z_1 . In principle, eq. (2.9) might be used to determine $G_w(z)$. In practice, no off-the-shelf technique is known to solve it, the main difficulty being that it relates $G_w(z)$ at different scales. In sect. 4, we obtain by other means an approximation to $\phi(w)$ which holds in the asymptotic regime $w \rightarrow \infty$ and reads

$$\phi(w) \approx \frac{1}{\zeta(\alpha - 1, n_{\min})} \frac{1}{w^\alpha}. \quad (2.10)$$

Dashed lines in Fig. 5 (top) represent eq. (2.10). The agreement with simulation data is reasonably good, yet our asymptotic estimate worsens as α and n_{\min} increase.

The variable of interest for our aims is the number of preferences V_t , introduced in eq. (2.2). From the above discussion, we conclude that V_t diverges as $t \rightarrow \infty$ with the same growth rate as Z_t . Therefore, it is convenient to introduce the variable

$$H_t = \frac{V_t}{m_Q^t} = W_t + \frac{1}{m_Q} W_{t-1} + \frac{1}{m_Q^2} W_{t-2} + \dots \quad (2.11)$$

As $t \rightarrow \infty$, the sequence $(H_t)_{t \geq 0}$ converges to a finite limit H . In Fig. 5 (bottom), we show the *p.d.f.* $\chi(h)$ of H obtained from the same numerical simulations described above. The distribution of H looks very similar to that of W and it is easy to understand why. First, we notice that

$$\mathbb{E}[H_t] = \mathbb{E}[W_t] + \frac{1}{m_Q} \mathbb{E}[W_{t-1}] + \frac{1}{m_Q^2} \mathbb{E}[W_{t-2}] + \dots = \frac{m_Q - 1/m_Q^t}{m_Q - 1}. \quad (2.12)$$

It follows that $\mathbb{E}[H] = m_Q/(m_Q - 1)$. Now, W_t, W_{t-1}, \dots converge to different copies of W as $t \rightarrow \infty$. However, these are not i.i.d. variables since they are correlated by construction (W_t is realized from W_{t-1} , which is realized from W_{t-2} , etc.). Of course, the correlation between W_t and W_{t-s} vanishes as s increases. Since $m_Q \sim \mathcal{O}(10)$, it follows that terms which are progressively less correlated with W_t are also quickly suppressed in eq. (2.11). Hence, it makes sense to approximate H by

$$H \simeq \frac{m_Q}{m_Q - 1} W, \quad (2.13)$$

and its *p.d.f.* $\chi(h)$ in the asymptotic regime $h \rightarrow \infty$ by

$$\chi(h) \approx \frac{h_0^{\alpha-1}}{\zeta(\alpha - 1, n_{\min})} \frac{1}{h^\alpha}, \quad (2.14)$$

⁴Supercritical Galton–Watson trees with power–law structure are studied in ref. [10] under the assumption that the offspring distribution fulfills $p(0) > 0$. As a consequence $\lim_{t \rightarrow \infty} \mathbb{P}\{Z_t = 0\} > 0$ in this case, *i.e.* the subspace of finite trees has a positive measure. The authors of ref. [10] are interested only in this subspace. Therefore their methodology does not apply here.

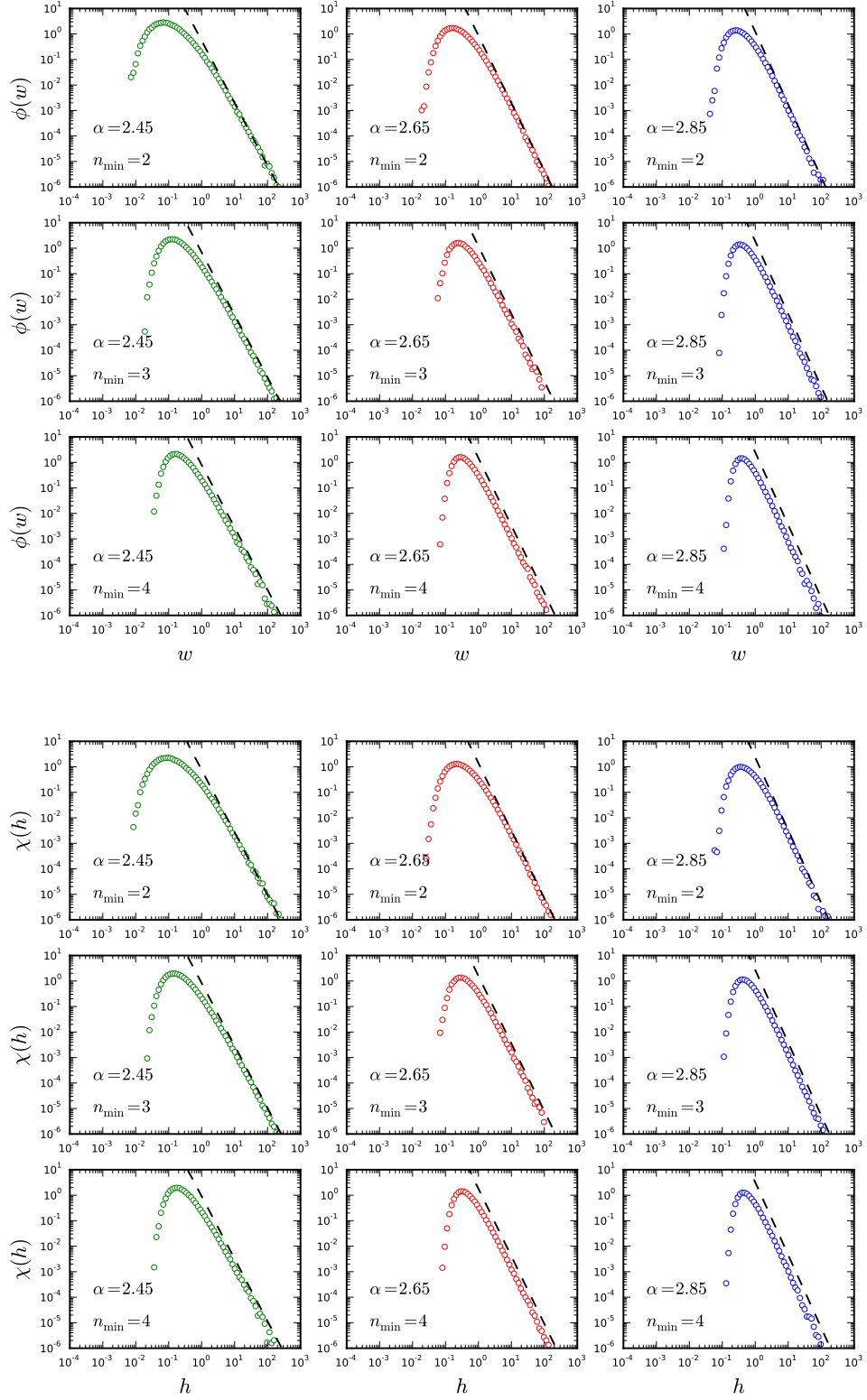


Figure 5 – (Top) *p.d.f.* of the limit variable W ; (Bottom) *p.d.f.* of the limit variable H . Dashed lines correspond to the analytic estimates reported resp. in eq. (2.10) and eq. (2.14).

$\mathbb{P}\{H < 1\}$		α			$\mathbb{P}\{H < 8\}$		α		
		2.45	2.65	2.85			2.45	2.65	2.85
	2	0.788	0.698	0.616		2	0.981	0.982	0.984
n_{\min}	3	0.796	0.724	0.660	n_{\min}	3	0.985	0.987	0.990
	4	0.798	0.734	0.683		4	0.986	0.989	0.992

Table 2 – Partition function $\mathcal{Z}_H(h) = \mathbb{P}\{H < h\}$ for $h = 1, 8$ and various choices of (α, n_{\min}) .

with $h_0 = m_Q/(m_Q - 1)$. Dashed lines in Fig. 5 (bottom) represent eq. (2.14). We observe that the quality of the approximation is analogous to that obtained for $\phi(w)$. We notice as well that $\chi(h)$ enters the asymptotic regime as $h \gtrsim 1$.

From Fig. 5 (bottom) we also see that a large fraction of the probability mass of H lies in the region $h \lesssim 1$. The partition function $\mathcal{Z}_H(h) = \mathbb{P}\{H < h\}$ saturates quickly to one as $h > 1$, as shown in Table 2, where values of $\mathcal{Z}_H(1)$ and $\mathcal{Z}_H(8)$ are reported by way of example for various choices of (α, n_{\min}) .

We conclude this section with two short remarks, that will be helpful in sect. 3. In first place, we let $p_t(v) = \mathbb{P}\{V_t = v\}$ denote the probability law of V_t and recall that V_t is an integer variable. Then, for $h = k/m_Q^t$ and $k \in \mathbb{N}$, we have

$$p_t(hm_Q^t) = \mathbb{P}\{V_t = hm_Q^t\} = \mathbb{P}\{H_t = h\} \rightarrow \chi(h)dh \quad \text{as } t \rightarrow \infty, \quad (2.15)$$

since H is a continuous variable. Secondly, the overall number of preferences on all trees is measured by the variable $\sum_{i=1}^Q V_t^{(i)}$, which diverges as $t \rightarrow \infty$ and also as $Q \rightarrow \infty$. For this reason, we find it convenient to introduce the variable

$$\bar{H}_t^{[Q]} \equiv \frac{1}{Q} \sum_{i=1}^Q H_t^{(i)}. \quad (2.16)$$

The sequence $(\bar{H}_t^{[Q]})_{t \geq 0}$ converges to a finite limit $\bar{H}^{[Q]}$ with continuous *p.d.f.* $\bar{\chi}^{[Q]}(h)$ as $t \rightarrow \infty$. We let $\bar{p}_t^{[Q]}(v) = \mathbb{P}\{\sum_{i=1}^Q V_t^{(i)} = v\}$ for $v \in \mathbb{N}$. Then, in perfect analogy with eq. (2.15), for $h = k/Qm_Q^t$ and $k \in \mathbb{N}$ we have $\bar{p}_t^{[Q]}(hQm_Q^t) = \mathbb{P}\{\sum_{i=1}^Q V_t^{(i)} = hQm_Q^t\} = \mathbb{P}\{\bar{H}_t^{[Q]} = h\} \rightarrow \bar{\chi}^{[Q]}(h)dh$ as $t \rightarrow \infty$. We shall say more about the distribution of $\bar{H}^{[Q]}$ in sect. 3.3.

3 Large-list limit in the quenched model

In sect. 2 we mentioned that numerical simulations of $F_Q(x|Q, N)$ under SR1 yield in most cases a tree $T^{(i)}$ for which the variable $Z_{t_{\text{exit}}}^{(i)}$ is only partially realized. This occurs whenever the generation of new vertices stops before those on the $(t_{\text{exit}} - 1)$ th level have fully branched. In such cases the overall number of vertices on $T^{(i)}$ is not correctly counted by the variable $V_{t_{\text{exit}}}^{(i)}$ introduced in eq. (2.2). In order to calculate $F_Q(x|Q, N)$, we either introduce *ad-hoc* variables that count subsets of vertices on the levels of a tree, or we reject *a priori* all Q -forests for which the above issue arises, thus obtaining an estimate of the distribution (the goodness of which we can only assess *a posteriori*). In this work we follow the latter approach. There are several possible approximations, lying between the following two extremes:

- the coarsest approximation consists in restricting the set of contributing Q -forests to those for which *all* of $(Z_{t_{\text{exit}}}^{(i)})_{i=1}^Q$ are fully realized when SR1 is fulfilled. We refer to this set $\mathcal{D}_{\min}(Q, N) \subset \mathfrak{T}$ as the *minimal ensemble* for SR1.

- the finest approximation consists in restricting the set of contributing Q -forests to those for which *each* of $(Z_{t_{\text{exit}}}^{(i)})_{i=1}^Q$ is either fully realized or not realized at all when SR1 is fulfilled. We refer to this set $\mathcal{D}_{\max}(Q, N) \subset \mathfrak{T}$ as the *maximal ensemble* for SR1.

Obviously, we have $\mathcal{D}_{\min}(Q, N) \subset \mathcal{D}_{\max}(Q, N)$, therefore we expect the estimate of $F_Q(x|Q, N)$ to be worse in $\mathcal{D}_{\min}(Q, N)$ than in $\mathcal{D}_{\max}(Q, N)$. Notice that the above issue does not arise under either SR2 or SR3. In this section, we first work out the large-list limit of $F_Q(x|Q, N)$ in the *minimal ensemble* for SR1. This case study allows us to set up the notation. Then, we calculate the vote distribution in the *maximal ensemble* for SR1 and finally under SR2 and SR3.

3.1 $F_Q(x|Q, N)$ in the *minimal ensemble* for SR1

As previously explained, we assume Q trees and stop the word-of-mouth spreading precisely when the overall number of vertices generated on all trees equals N . Accordingly, the space of all possible vote configurations is given by the discrete simplex

$$\Sigma(Q, N) = \left\{ (v_1, \dots, v_Q) \in \mathbb{N}^Q : \sum_{k=1}^Q v_k = N \right\}. \quad (3.1)$$

Given $v = (v_1, \dots, v_Q) \in \Sigma(Q, N)$, the probability that at time t the vote variables defined in eq. (2.2) fulfill $\{V_t^{(1)} = v_1, \dots, V_t^{(Q)} = v_Q\}$ amounts to

$$\Pi_t(v|Q, N) = \frac{p_t(v_1) \cdots p_t(v_Q)}{\sum_{v \in \Sigma(Q, N)} p_t(v_1) \cdots p_t(v_Q)}. \quad (3.2)$$

It is important to recognize that $\Pi_t(\cdot|Q, N)$ measures subsets of \mathfrak{T} . More precisely, the domain of $\Pi_t(\cdot|Q, N)$ is the *ensemble* $\mathcal{D}_t(Q, N) \subset \mathfrak{T}$ defined by

$$\mathcal{D}_t(Q, N) = \left\{ T \in \mathfrak{T} : (V_t(T^{(1)}), \dots, V_t(T^{(Q)})) \in \Sigma(Q, N) \right\}. \quad (3.3)$$

Sampling $\Pi_t(\cdot|Q, N)$ is a hard job. In order to generate an acceptable sample $T \in \mathcal{D}_t(Q, N)$ in a computer simulation, we are supposed to run eq. (2.4) up to the t th level (which means drawing *all* variables $(Z_s^{(i)})_{i=1}^Q$ for $s \leq t$) and then check that $(V_t(T^{(1)}), \dots, V_t(T^{(Q)})) \in \Sigma(Q, N)$. Needless to say, such samples have an extremely low probability of being generated. Starting from eq. (3.3), we define

$$\mathcal{D}_{\min}(Q, N) = \bigcup_{t=0}^{\infty} \mathcal{D}_t(Q, N). \quad (3.4)$$

This is the *minimal ensemble* for SR1. For all $T \in \mathcal{D}_{\min}(Q, N)$, SR1 is exactly fulfilled at some level t_{exit} with all of $(Z_{t_{\text{exit}}}^{(i)})_{i=1}^Q$ being fully counted. We call it *minimal* since it is the smallest subset of \mathfrak{T} in which $F_Q(x|Q, N)$ under SR1 can be consistently represented as a functional of the probabilities of $\{(V_t^{(i)})_{i=1}^Q\}_{t \in \mathbb{N}}$ alone. Of course, the majority of $T \in \mathfrak{T}$ contributing to numerical simulations of $F_Q(x|Q, N)$ under SR1 lie outside $\mathcal{D}_{\min}(Q, N)$.

Notice that since $\mathcal{D}_t(Q, N) \cap \mathcal{D}_{t'}(Q, N) = \emptyset$ for $t' \neq t$, it follows that

$$\mathbb{P}\{T \in \mathcal{D}_{\min}(Q, N)\} = \sum_{t=0}^{\infty} \mathbb{P}\{T \in \mathcal{D}_t(Q, N)\}. \quad (3.5)$$

The above probabilities come into play when observing that given (Q, N) , the actual level t_{exit} at which the algorithm stops is unknown. For this reason we introduce the stopping time

$$\tau_{Q, N} = \min \left\{ t \in \mathbb{N} : \sum_{i=1}^Q V_t^{(i)} \geq N \right\}, \quad (3.6)$$

representing the earliest level at which the sum of the vote variables defined in eq. (2.2) exceeds N . Similar to $V_t^{(i)}$, also $\tau_{Q,N}$ is a stochastic variable, *i.e.* it is a function of $T \in \mathfrak{T}$. As such, it is differently realized in each iteration of the algorithm. Of course, $\tau_{Q,N}$ fulfills

$$\mathbb{P}\{\tau = t \mid \mathcal{D}_{\min}\} \equiv \mathbb{P}\{\tau_{Q,N}(T) = t \mid T \in \mathcal{D}_{\min}(Q, N)\} = \frac{\mathbb{P}\{T \in \mathcal{D}_t(Q, N)\}}{\mathbb{P}\{T \in \mathcal{D}_{\min}(Q, N)\}}. \quad (3.7)$$

Sampling $\mathbb{P}\{\tau = t \mid \mathcal{D}_{\min}\}$ is just as difficult as sampling $\Pi_t(\cdot \mid Q, N)$: on the one hand $\mathbb{P}\{T \in \mathcal{D}_{\min}\} \rightarrow 0$ as $Q, N \rightarrow \infty$, on the other hand we do not know any algorithmic recipe to *a priori* generate Q -forests $T \in \mathcal{D}_{\min}(Q, N)$ with correct probability. However, we can obtain an estimate of $\mathbb{P}\{\tau = t \mid \mathcal{D}_{\min}\}$ by measuring the frequency of those $T \in \mathfrak{T}$ for which SR1 is fulfilled with either $Z_t^{(Q)}$ or $Z_{t+1}^{(1)}$ being partially counted.

Now, given (Q, N) and $x = \ell Q/N$ with $\ell = 1, 2, \dots, N$, the *discrete* probability that the excess-of-votes variable yields x in the *minimal ensemble* for SR1 is given by

$$F_Q^{(i)}(x|Q, N; \mathcal{D}_{\min}) = \sum_{t=0}^{\infty} \left[\sum_{v \in \Sigma(Q, N)} \delta_{xN/Q, v_i} \Pi_t(v|Q, N) \right] \mathbb{P}\{\tau = t \mid \mathcal{D}_{\min}\}, \quad (3.8)$$

with the δ -symbol within square brackets representing the Kronecker delta. Although we need to select one specific candidate to count preferences, the upper index (i) on the l.h.s. is redundant, since $\Pi_t(v|Q, N)$ is invariant under permutations of the components of the vote vector v . Hence, we drop it in the sequel. We wish to work out eq. (3.8) in the large-list limit, *i.e.* as $Q, N \rightarrow \infty$. To this aim, we notice that the simplex constraint can be represented in terms of another δ -symbol: for any function $f(v)$ the identity

$$\sum_{v \in \Sigma(Q, N)} f(v) = \sum_{v_1, \dots, v_Q=1}^M \delta_{\sum_k v_k, N} f(v) = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\lambda \sum_{v_1, \dots, v_Q=1}^M e^{i\lambda(\sum_k v_k - N)} f(v_1, \dots, v_Q) \quad (3.9)$$

holds for all integers $M \geq N$. In particular, the rightmost equality follows from representing the Kronecker delta in Fourier harmonics. Upon using eq. (3.9) and the explicit expression of $\Pi_t(v|Q, N)$, we turn eq. (3.8) into

$$F_Q(x|Q, N; \mathcal{D}_{\min}) = \sum_{t=0}^{\infty} p_t \left(\frac{xN}{Q} \right) \left[\frac{\int_{-\pi}^{\pi} d\lambda e^{-i\lambda \frac{N}{Q}(Q-x)} \prod_{j \neq i}^{1 \dots Q} \sum_{v_j=1}^M e^{i\lambda v_j} p_t(v_j)}{\int_{-\pi}^{\pi} d\lambda e^{-i\lambda N} \prod_{j=1}^Q \sum_{v_j=1}^M e^{i\lambda v_j} p_t(v_j)} \right] \mathbb{P}\{\tau = t \mid \mathcal{D}_{\min}\}. \quad (3.10)$$

Provided $x < Q$, the exponential $\exp\{-i\lambda \frac{N}{Q}(Q-x)\}$ at numerator oscillates quickly as $N \rightarrow \infty$ and thus lets the integral over λ receive contributions only from a region around $\lambda = 0$ with size proportional to $\pi Q/[N(Q-x)] \rightarrow 0$. Analogously, the exponential $\exp\{-i\lambda N\}$ at denominator oscillates quickly as $N \rightarrow \infty$ and thus lets the integral over λ receive contributions only from a region around $\lambda = 0$ with size proportional to $\pi/N \rightarrow 0^5$. Therefore, the ratio in square brackets converges quickly to $Q/[(Q-x) \sum_{v_i=0}^M p_t(v_i)]$ as $N \rightarrow \infty$. Clearly, in this approximation $F_Q(x|Q, N)$ diverges as $x \rightarrow Q$, the reason being that the original integral at numerator stops oscillating in that limit and the above argument fails. However, we can approximate the diverging factor $Q/(Q-x) \rightarrow 1$ as $Q \rightarrow \infty$. We finally set $M = N$. This yields

$$F_Q(x|Q, N; \mathcal{D}_{\min}) = \sum_{t=0}^{\infty} \frac{p_t(xN/Q)}{\sum_{v=1}^N p_t(v)} \mathbb{P}\{\tau = t \mid \mathcal{D}_{\min}\}, \quad \text{as } Q, N \rightarrow \infty. \quad (3.11)$$

⁵To make this point clear: we extend $F_Q(x|Q, N)$ to $N \in \mathbb{C}$ by analytic continuation, then we rotate $N \rightarrow -i\tilde{N}$. This turns the oscillating integrals into exponentially damped ones. Everything under the integral sign (except the highly oscillating exponentials) can be then evaluated at $\lambda = 0$ and be taken out of the integrals. At the end we rotate back $\tilde{N} \rightarrow iN$. It is funny to notice that highly oscillating integrals such as eq. (3.10) arise in applicative contexts which have nothing to do with opinion dynamics. An example is represented by the Bjorken scaling in QCD (see for instance eq. (9.5) of ref. [11]), which is analogous in some respect to the FC scaling.

To highlight the transition to a continuous distribution, we find it convenient to set $N \rightarrow N_s = h_0 Q m_Q^s$ (recall that $\mathbb{E}[V_t] = h_0 m_Q^t$) and let it diverge along the sequence $s = 1, 2, 3, \dots, \infty$ (thermodynamic limit). If N_s is not an integer, we define $F_Q(x|Q, N_s; \mathcal{D}_{\min})$ by linear interpolation between its values at the nearest integers $N_- = \lfloor N_s \rfloor$ and $N_+ = \lceil N_s \rceil$. This setting is legitimate as well as reasonable: on the one hand we know that preferences distribute symmetrically among the Q trees up to fluctuations, on the other hand the variable $s = \log[N/(h_0 Q)]/\log m_Q$ just represents the average level at which the word of mouth stops given (Q, N) . In other words, the probability distribution $\mathbb{P}\{\tau = t | \mathcal{D}_{\min}\}$ as a function of t peaks precisely at $t \simeq s$ for $N = N_s$. It follows from eq. (2.15) that $p_t(xN/Q) = p_t(h_0 x m_Q^s) = p_t(h_0 x m_Q^{s-t} m_Q^t) \rightarrow h_0 m_Q^{s-t} \chi(h_0 x m_Q^{s-t}) dx$ as $s \rightarrow \infty$. In the same limit the probability distribution of x turns into a continuous distribution, namely $F_Q(x|Q, N_s) \rightarrow \mathcal{F}_Q(x|Q, N_s) dx$. We conclude that

$$\mathcal{F}_Q(x|Q, N_s; \mathcal{D}_{\min}) = \sum_{t=s, s\pm 1, s\pm 2, \dots} \frac{h_0 m_Q^{s-t} \chi(h_0 x m_Q^{s-t})}{\int_0^{h_0 Q m_Q^{s-t}} dy \chi(y)} \mathbb{P}\{\tau = t | \mathcal{D}_{\min}\}, \quad \text{as } Q, s \rightarrow \infty. \quad (3.12)$$

The above distribution is correctly normalized, as can be seen upon integrating both sides over $x \in (0, Q)$. Provided $h_0 Q \gtrsim m_Q^{t-s+1}$, the integral at denominator can be omitted without significant loss of accuracy (see Table 2).

Remark 3.1. While eq. (3.8) is exact, eq. (3.12) is formally correct provided $x \ll Q$. An important difference between the two distributions arises as we calculate their expectations, which should equal one by symmetry. Indeed, we have

$$\begin{aligned} \mathbb{E}[x|Q, N; \mathcal{D}_{\min}] &= \frac{1}{Q} \sum_{i=1}^Q \sum_x x F_Q^{(i)}(x|Q, N; \mathcal{D}_{\min}) \\ &= \frac{1}{N} \sum_{t=0}^{\infty} \left[\sum_{v \in \Sigma(Q, N)} \sum_{i=1}^Q v_i \Pi_t(v|Q, N) \right] \mathbb{P}\{\tau = t | \mathcal{D}_{\min}\} = \sum_{t=0}^{\infty} \mathbb{P}\{\tau = t | \mathcal{D}_{\min}\} = 1, \end{aligned} \quad (3.13)$$

whereas

$$\begin{aligned} \mathcal{E}[x|Q, N_s; \mathcal{D}_{\min}] &= \int_0^Q dx x \mathcal{F}_Q(x|Q, N_s; \mathcal{D}_{\min}) \\ &= \sum_{t=s, s\pm 1, s\pm 2, \dots} \frac{m^{t-s}}{h_0} \frac{\int_0^{h_0 Q m_Q^{s-t}} dy y \chi(y)}{\int_0^{h_0 Q m_Q^{s-t}} dy \chi(y)} \mathbb{P}\{\tau = t | \mathcal{D}_{\min}\}. \end{aligned} \quad (3.14)$$

For finite values of Q we have in general $\mathcal{E}[x|Q, N_s; \mathcal{D}_{\min}] \neq 1$. Notice that the ratio of integrals on the r.h.s. of eq. (3.14) converges to h_0 as $Q \rightarrow \infty$. Hence, in this limit we are left with $\mathcal{E}[x|Q, N_s; \mathcal{D}_{\min}] \simeq \sum_t m^{t-s} \mathbb{P}\{\tau = t | \mathcal{D}_{\min}\}$. The value of the latter sum depends on how τ_{Q, N_s} distributes around s . In full generality we must regard $\mathcal{E}[x|Q, N_s; \mathcal{D}_{\min}]$ as a measurement of the quality by which eq. (3.12) approximates $F_Q(x|Q, N; \mathcal{D}_{\min})$ in the large-list limit. ■

In Table 3 (left) we report an estimate of the probability $\mathbb{P}\{\tau = t | \mathcal{D}_{\min}\}$ for $s = 5$, $(\alpha, n_{\min}) = (2.45, 3)$ and $Q = 8, \dots, 256$. We obtained this estimate from numerical simulations just as explained above, *i.e.* by measuring the frequency of those $T \in \mathfrak{T}$ for which SR1 is fulfilled with either $Z_t^{(Q)}$ or $Z_{t+1}^{(1)}$ being partially counted. In Table 3 (right) we report an estimate of the expectation $\mathcal{E}[x|Q, N_s; \mathcal{D}_{\min}]$ corresponding to our estimate of $\mathbb{P}\{\tau = t | \mathcal{D}_{\min}\}$ for the same choice of parameters. Analogous simulations at $s = 6$ display a rigid shift of the probability distribution of τ_{Q, N_s} around $t = 6$ with no evident change in the structure of the tails. In full generality we conclude that τ_{Q, N_s} peaks at $t = s$ with increasing probability as $Q \rightarrow \infty$, provided s is sufficiently large. By extrapolation we have

$$\lim_{Q \rightarrow \infty} \lim_{s \rightarrow \infty} \mathbb{P}\{\tau = t | \mathcal{D}_{\min}\} = \delta_{t,s}. \quad (3.15)$$

$\mathbb{P}\{\tau = t \mid \mathcal{D}_{\min}\}$		t				Q	$\mathcal{E}[x \mid Q, N_5; \mathcal{D}_{\min}]$	$\mathcal{C}(Q)$
		3	4	5	6			
Q	8	0.0018(2)	0.0447(8)	0.9524(8)	0.0009(1)	8	1.18(4)	0.97
	16	0.0010(2)	0.0266(7)	0.9723(7)	n/a	16	1.10(3)	0.97
	32	0.0005(1)	0.0157(8)	0.9837(8)	n/a	32	1.05(3)	0.98
	64	0.0002(1)	0.0082(5)	0.9915(5)	n/a	64	1.03(3)	0.99
	128	n/a	0.0048(7)	0.9951(7)	n/a	128	1.02(3)	0.99
	256	n/a	0.0023(6)	0.9973(5)	n/a	256	1.00(3)	0.99

Table 3 – (Left) Approximate probability of τ_{Q, N_s} in \mathcal{D}_{\min} for $s = 5$, $t = 3, \dots, 6$, $(\alpha, n_{\min}) = (2.45, 3)$ and $Q = 8, \dots, 256$. Bootstrap uncertainty is quoted in parentheses. N/a entries correspond to measurements for which noise/signal ≥ 1 . (Right) Expectation of x in the large-list limit and $\mathcal{C}(Q)$ corresponding to our estimate of $\mathbb{P}\{\tau = t \mid \mathcal{D}_{\min}\}$.

As a consequence, $\mathcal{E}[x \mid Q, N_s; \mathcal{D}_{\min}] \rightarrow 1$ as $Q, s \rightarrow \infty$. From Table 3 we notice that $\mathcal{E}[x \mid Q, N_s; \mathcal{D}_{\min}]$ deviates from one by less than 20% even at the smallest value of Q among those simulated.

We now go back to eq. (3.12). From eq. (3.15) we infer that

$$\mathcal{F}_Q^*(x) = \lim_{Q \rightarrow \infty} \lim_{s \rightarrow \infty} \mathcal{F}_Q(x \mid Q, N_s; \mathcal{D}_{\min}) = h_0 \chi(h_0 x). \quad (3.16)$$

Thus we see that $\mathcal{F}_Q^*(x)$ has an asymptotic power-law behaviour. The asymptotic regime is reached as soon as $x \gtrsim h_0^{-1}$. Eq. (2.14) yields a prediction of the exponent and the scale coefficient of the power law in the large-list limit. If we forget about the structure of $\mathbb{P}\{\tau = t \mid \mathcal{D}_{\min}\}$ for a while and concentrate on the general structure eq. (3.12), we see that in the large-list limit $\mathcal{F}_Q(x \mid Q, N_s; \mathcal{D}_{\min})$ is given by a convolution of χ at largely separated scales (recall that $m_Q \sim \mathcal{O}(10)$). Given x , we can split the latter into two groups, namely forward ($m_Q x, m_Q^2 x, \dots$) and backward ($x/m_Q, x/m_Q^2, \dots$) scales. Explicitly, we have

$$\begin{aligned} \mathcal{F}_Q(x \mid Q, N_s; \mathcal{D}_{\min}) &= \frac{h_0 \chi(h_0 x)}{\int_0^{h_0 Q} dy \chi(y)} \mathbb{P}\{\tau = s \mid \mathcal{D}_{\min}\} \\ &+ m_Q \frac{h_0 \chi(h_0 m_Q x)}{\int_0^{h_0 Q m_Q} dy \chi(y)} \mathbb{P}\{\tau = s - 1 \mid \mathcal{D}_{\min}\} + m_Q^2 \frac{h_0 \chi(h_0 m_Q^2 x)}{\int_0^{h_0 Q m_Q^2} dy \chi(y)} \mathbb{P}\{\tau = s - 2 \mid \mathcal{D}_{\min}\} + \dots \quad (\text{fw. terms}) \\ &+ \frac{1}{m_Q} \frac{h_0 \chi(h_0 x/m_Q)}{\int_0^{h_0 Q/m_Q} dy \chi(y)} \mathbb{P}\{\tau = s + 1 \mid \mathcal{D}_{\min}\} + \frac{1}{m_Q^2} \frac{h_0 \chi(h_0 x/m_Q^2)}{\int_0^{h_0 Q/m_Q^2} dy \chi(y)} \mathbb{P}\{\tau = s + 2 \mid \mathcal{D}_{\min}\} + \dots \quad (\text{bw. terms}) \end{aligned} \quad (3.17)$$

If $\chi(h_0 x)$ is in the power-law regime, this is even more the case for $\chi(h_0 m_Q^k x)$, $k = 1, 2, \dots$. We thus see that forward terms are safe: they do not spoil the overall power-law behaviour, do not even change the exponent of the power law and only contribute by modifying its scale coefficient. By contrast, backward terms are potentially dangerous: they shift x towards regions where χ is not anymore in the asymptotic regime. Backward terms are in principle able to impair the power-law behaviour. Nevertheless, they are suppressed by inverse powers of m_Q . This suggests that power-law structures in the scaling vote distribution may emerge even at moderate values of Q . Keeping only terms corresponding to $\tau_{Q, N_s} = s, s \pm 1$ and assuming that $x/m_Q \gtrsim 1$ yields the asymptotic limit

$$\begin{aligned} \mathcal{F}_Q(x \mid Q, N_s; \mathcal{D}_{\min}) &\simeq \frac{\mathcal{C}(Q)}{\zeta(\alpha - 1, n_{\min})} \frac{1}{x^\alpha}, \quad \text{as } x, Q, s \rightarrow \infty, \quad \text{with} \\ \mathcal{C}(Q) &= \frac{\mathbb{P}\{\tau = s \mid \mathcal{D}_{\min}\}}{\int_0^{h_0 Q} dy \chi(y)} + \frac{1}{m_Q^{\alpha-1}} \frac{\mathbb{P}\{\tau = s - 1 \mid \mathcal{D}_{\min}\}}{\int_0^{h_0 Q m_Q} dy \chi(y)} + m_Q^{\alpha-1} \frac{\mathbb{P}\{\tau = s + 1 \mid \mathcal{D}_{\min}\}}{\int_0^{h_0 Q/m_Q} dy \chi(y)}. \end{aligned} \quad (3.18)$$

The coefficient $\mathcal{C}(Q)$ represents the only correction to the power-law scaling at finite Q . It fulfills $\mathcal{C}(Q) \simeq 1$ and $\lim_{Q \rightarrow \infty} \mathcal{C}(Q) = 1$. For $Q \gg m_Q$ the integrals at denominator can be omitted, as previously noticed. In the specific case of the *minimal ensemble*, we can set $\mathcal{C}(Q) = 1$ for all Q without loss of accuracy, as the values reported in Table 3 (right) suggest.

3.2 $F_Q(x|Q, N)$ in the *maximal ensemble* for SR1

Now that we have shown in a simplified framework how to calculate $F_Q(x|Q, N)$ in the large-list limit under SR1, we proceed to work out a much better estimate of it. As mentioned at the beginning of the section, the finest approximation in terms of level variables takes into account all Q -forests for which *each* of $(Z_{t_{\text{exit}}}^{(i)})_{i=1}^Q$ is either fully realized or not realized at all when SR1 is fulfilled. To formalize this, we go through the same steps as we did in last section. Concretely, given (Q, N) we introduce the *ensembles*

$$\mathcal{D}_{t,q}(Q, N) = \left\{ T \in \mathfrak{T} : \left(V_{t+1}(T^{(1)}), \dots, V_{t+1}(T^{(q)}), V_t(T^{(q+1)}), \dots, V_t(T^{(Q)}) \right) \in \Sigma(Q, N) \right\}, \quad (3.19)$$

for $t \geq 0$ and $1 \leq q \leq Q$. Notice that $\mathcal{D}_{t,q}(Q, N) = \mathcal{D}_{t+1}(Q, N)$. Moreover, independently of t or q we expect that $\mathbb{P}\{T \in \mathcal{D}_{t,q}(Q, N)\} \ll 1$, since $\mathcal{D}_{t,q}(Q, N)$ is a highly restricted subset of \mathfrak{T} . Starting from eq. (3.19) we define

$$\mathcal{D}_{\text{max}}(Q, N) = \bigcup_{t=0}^{\infty} \bigcup_{q=1}^Q \mathcal{D}_{t,q}(Q, N). \quad (3.20)$$

This is the *maximal ensemble* for SR1. For all $T \in \mathcal{D}_{\text{max}}(Q, N)$, SR1 is exactly fulfilled at some level t_{exit} with each of $(Z_{t_{\text{exit}}}^{(i)})_{i=1}^Q$ being either fully counted or not counted at all. We call it *maximal* since it is the largest subset of \mathfrak{T} in which $F_Q(x|Q, N)$ under SR1 can be consistently represented as a functional of the probabilities of $\{(V_t^{(i)})_{i=1}^Q\}_{t \in \mathbb{N}}$ alone.

Given $v \in \Sigma(Q, N)$, the probability that the vote variables fulfill $\{V_{t+1}^{(1)} = v_1, \dots, V_{t+1}^{(q)} = v_q, V_t^{(q+1)} = v_{q+1}, \dots, V_t^{(Q)} = v_Q\}$ amounts to

$$\Pi_{t,q}(v|Q, N) = \frac{p_{t+1}(v_1) \cdots p_{t+1}(v_q) p_t(v_{q+1}) \cdots p_t(v_Q)}{\sum_{v \in \Sigma(Q, N)} p_{t+1}(v_1) \cdots p_{t+1}(v_q) p_t(v_{q+1}) \cdots p_t(v_Q)}. \quad (3.21)$$

The domain of $\Pi_{t,q}(v|Q, N)$ is precisely $\mathcal{D}_{t,q}(Q, N)$. Notice that since $\mathcal{D}_{t,q} \cap \mathcal{D}_{t',q'} = \emptyset$ for $(t, q) \neq (t', q')$, it follows that

$$\mathbb{P}\{T \in \mathcal{D}_{\text{max}}(Q, N)\} = \sum_{t=0}^{\infty} \sum_{q=1}^Q \mathbb{P}\{T \in \mathcal{D}_{t,q}(Q, N)\}. \quad (3.22)$$

The above probabilities come into play when observing that the actual values $(t_{\text{exit}}, q_{\text{exit}})$ at which the algorithm stops fluctuate randomly along different iterations. For this reason we introduce a fractional stopping time

$$\tau_{Q,N}^{\text{frac}} = \inf \left\{ t + \frac{q}{Q} : \sum_{i=1}^q V_{t+1}^{(i)} + \sum_{i=q+1}^Q V_t^{(i)} \geq N \right\}. \quad (3.23)$$

While $\tau_{Q,N}^{\text{frac}}$ is a discrete variable for finite values of Q , it becomes continuous as $Q \rightarrow \infty$. For finite Q the “inf” is actually a “min” and since $1 \leq q \leq Q$, there exist no $(t_1, q_1), (t_2, q_2)$ with $(t_1, q_1) \neq (t_2, q_2)$ such that $\tau_{Q,N}^{\text{frac}} = t_1 + q_1/Q = t_2 + q_2/Q$. The probability distribution of $\tau_{Q,N}^{\text{frac}}$ is given by

$$\mathbb{P} \left\{ \tau_{Q,N}^{\text{frac}} = t + \frac{q}{Q} \mid \mathcal{D}_{\text{max}} \right\} \equiv \mathbb{P} \left\{ \tau_{Q,N}^{\text{frac}}(T) = t + \frac{q}{Q} \mid T \in \mathcal{D}_{\text{max}}(Q, N) \right\} = \frac{\mathbb{P}\{T \in \mathcal{D}_{t,q}(Q, N)\}}{\mathbb{P}\{T \in \mathcal{D}_{\text{max}}(Q, N)\}}. \quad (3.24)$$

$\mathbb{P}_{\text{int}}\{\tau = t \mathcal{D}_{\text{max}}\}$		t				Q	$\mathcal{E}[x Q, N_5; \mathcal{D}_{\text{max}}]$	$\mathcal{C}(Q)$
		3	4	5	6			
Q	8	0.0031(1)	0.1194(5)	0.7787(6)	0.0987(3)	8	0.81(4)	2.79
	16	0.0023(1)	0.1093(5)	0.8126(4)	0.0756(2)	16	0.91(3)	1.82
	32	0.0015(1)	0.1012(4)	0.8418(4)	0.0554(1)	32	0.97(3)	1.84
	64	0.0012(1)	0.0940(5)	0.8650(4)	0.0397(1)	64	1.00(3)	0.87
	128	0.0010(1)	0.0863(4)	0.8850(4)	0.0277(1)	128	1.00(3)	0.88
	256	0.0007(1)	0.0778(5)	0.9029(4)	0.0187(1)	256	0.97(3)	0.90

Table 4 – (Left) Approximate integrated probability of $\tau_{Q,N_s}^{\text{frac}}$ in \mathcal{D}_{max} for $s = 5$, $t = 3, \dots, 6$, $(\alpha, n_{\text{min}}) = (2.45, 3)$ and $Q = 8, \dots, 256$. Bootstrap uncertainty is quoted in parentheses; (Right) Expectation of x in the large-list limit and $\mathcal{C}(Q)$ corresponding to our estimate of $\mathbb{P}_{\text{int}}\{\tau = t | \mathcal{D}_{\text{max}}\}$.

Although sampling $\mathbb{P}\{\tau_{Q,N}^{\text{frac}} = t + q/Q | \mathcal{D}_{\text{max}}\}$ is very difficult (for the same reasons as previously), we can obtain an estimate of it by measuring the frequency of those $T \in \mathfrak{T}$ for which SR1 is fulfilled with $Z_{t+1}^{(q+1)}$ being partially counted.

Now, given (Q, N) and $x = \ell Q/N$ with $\ell = 1, 2, \dots, N$, the discrete probability that the excess-of-votes variable yields x in the *maximal ensemble* for SR1 is given by

$$F_Q(x|Q, N; \mathcal{D}_{\text{max}}) = \frac{1}{Q} \sum_{i=1}^Q \sum_{t=0}^{\infty} \sum_{q=1}^Q \left[\sum_{v \in \Sigma(Q, N)} \delta_{xN/Q, v_i} \Pi_{t,q}(v|Q, N) \right] \mathbb{P}\left\{\tau_{Q,N}^{\text{frac}} = t + \frac{q}{Q} \middle| \mathcal{D}_{\text{max}}\right\}. \quad (3.25)$$

By the same arguments used in sect. 3.1 we can show that the above expression boils down to

$$F_Q(x|Q, N; \mathcal{D}_{\text{max}}) = \sum_{t=0}^{\infty} \sum_{q=1}^Q \left[\frac{q}{Q} \frac{p_{t+1}(xN/Q)}{\sum_{v=1}^N p_{t+1}(v)} + \frac{Q-q}{Q} \frac{p_t(xN/Q)}{\sum_{v=1}^N p_t(v)} \right] \mathbb{P}\left\{\tau_{Q,N}^{\text{frac}} = t + \frac{q}{Q} \middle| \mathcal{D}_{\text{max}}\right\}, \quad (3.26)$$

as $Q, N \rightarrow \infty$. Each term on the r.h.s. combines linearly the probabilities of the vote variables V_t and V_{t+1} for some t . We can rearrange the sum so as to avoid such mixing. This yields

$$F_Q(x|Q, N; \mathcal{D}_{\text{max}}) = \sum_{t=0}^{\infty} \frac{p_t(xN/Q)}{\sum_{v=1}^N p_t(v)} \mathbb{P}_{\text{int}}\{\tau = t | \mathcal{D}_{\text{max}}\}, \quad (3.27)$$

with the integrated probability $\mathbb{P}_{\text{int}}\{\tau = t | \mathcal{D}_{\text{max}}\}$ being defined by

$$\mathbb{P}_{\text{int}}\{\tau = t | \mathcal{D}_{\text{max}}\} \equiv \frac{1}{Q} \sum_{q=1}^Q \left[q \mathbb{P}\left\{\tau_{Q,N}^{\text{frac}} = t - 1 + \frac{q}{Q} \middle| \mathcal{D}_{\text{max}}\right\} + (Q - q) \mathbb{P}\left\{\tau_{Q,N}^{\text{frac}} = t + \frac{q}{Q} \middle| \mathcal{D}_{\text{max}}\right\} \right]. \quad (3.28)$$

We see that eq. (3.27) differs from eq. (3.11) only in that $\mathbb{P}\{\tau = t | \mathcal{D}_{\text{min}}\}$ is replaced by $\mathbb{P}_{\text{int}}\{\tau = t | \mathcal{D}_{\text{max}}\}$. We also observe that $\mathbb{P}_{\text{int}}\{\tau = t | \mathcal{D}_{\text{max}}\}$ is a convex average of the probabilities of $\tau_{Q,N}^{\text{frac}}$ around t . The first term in square brackets on the r.h.s. of eq. (3.28) has maximal weight for $q = Q$ and for this value of q it yields $Q \mathbb{P}\{\tau_{Q,N}^{\text{frac}} = t | \mathcal{D}_{\text{max}}\}$. The second term in square bracket has maximal weight for $q = 1$ and for this value of q it yields $(Q - 1) \mathbb{P}\{\tau_{Q,N}^{\text{frac}} = t + 1/Q | \mathcal{D}_{\text{max}}\} \simeq Q \mathbb{P}\{\tau_{Q,N}^{\text{frac}} = t | \mathcal{D}_{\text{max}}\}$ as $Q \rightarrow \infty$. The other terms of eq. (3.28) are weighted by increasingly lower coefficients and yield the probability of the stopping time at fractional points which are increasingly far from t within the interval $(t - 1, t + 1)$. It can be easily checked that $\sum_{t=0}^{\infty} \mathbb{P}_{\text{int}}\{\tau = t | \mathcal{D}_{\text{max}}\} = 1$.

Upon setting $N \rightarrow N_s = h_0 Q m_Q^s$, we obtain our final estimate

$$\mathcal{F}_Q(x|Q, N_s; \mathcal{D}_{\text{max}}) = \sum_{t=s, s \pm 1, s \pm 2, \dots} \frac{h_0 m_Q^{s-t} \chi(h_0 x m_Q^{s-t})}{\int_0^{h_0 Q m_Q^{s-t}} dy \chi(y)} \mathbb{P}_{\text{int}}\{\tau = t | \mathcal{D}_{\text{max}}\}, \quad \text{as } Q, s \rightarrow \infty, \quad (3.29)$$

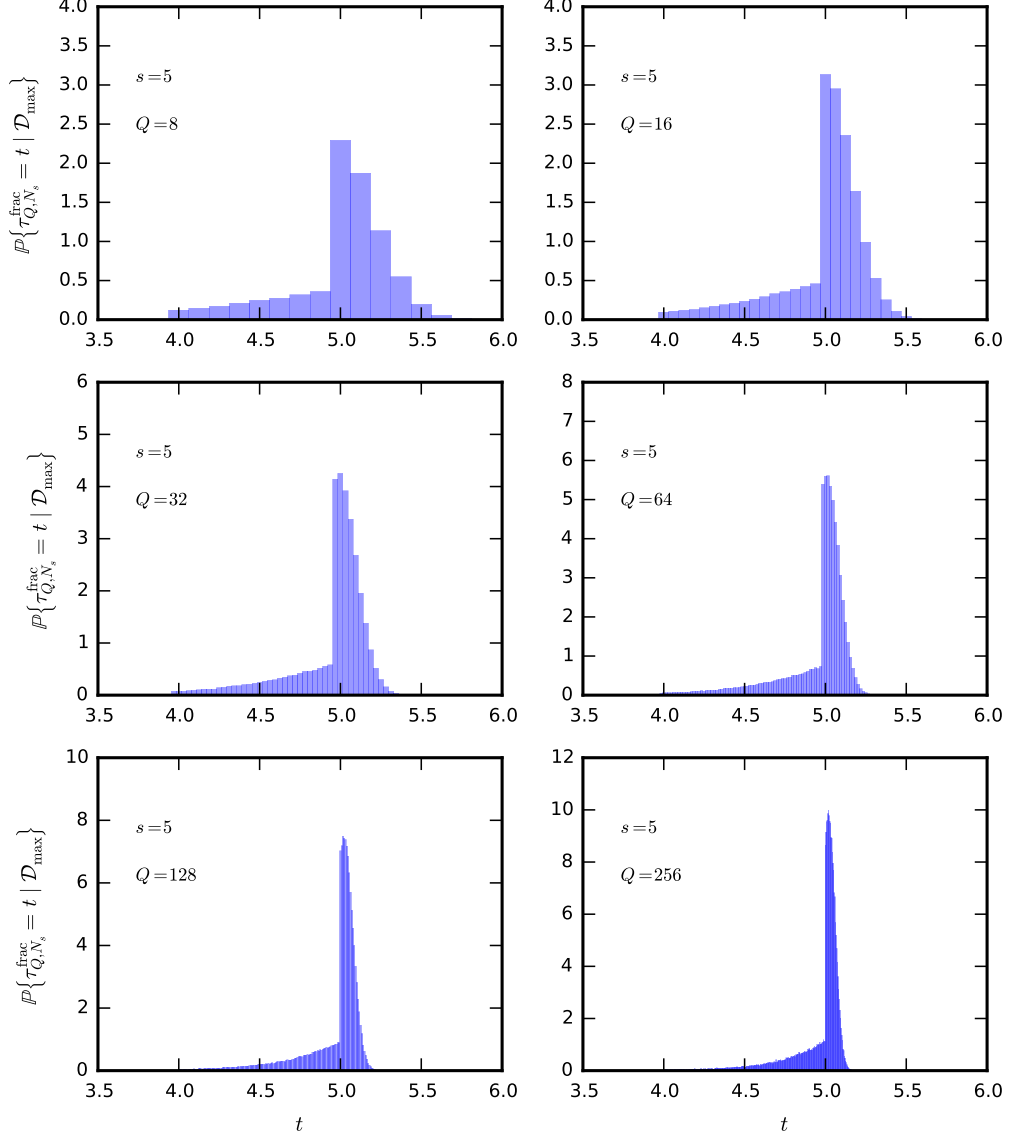


Figure 6 – Distribution of $\tau_{Q,N_s}^{\text{frac}}$ for $s = 5$, $(\alpha, n_{\min}) = (2.45, 3)$ and $Q = 8, \dots, 256$.

and its asymptotic limit

$$\mathcal{F}_Q(x|Q, N_s; \mathcal{D}_{\max}) \simeq \frac{\mathcal{C}(Q)}{\zeta(\alpha - 1, n_{\min})} \frac{1}{x^\alpha}, \quad \text{as } x, Q, s \rightarrow \infty, \quad \text{with}$$

$$\mathcal{C}(Q) = \frac{\mathbb{P}_{\text{int}}\{\tau = s | \mathcal{D}_{\max}\}}{\int_0^{h_0 Q} dy \chi(y)} + \frac{1}{m_Q^{\alpha-1}} \frac{\mathbb{P}_{\text{int}}\{\tau = s - 1 | \mathcal{D}_{\max}\}}{\int_0^{h_0 Q m_Q} dy \chi(y)} + m_Q^{\alpha-1} \frac{\mathbb{P}_{\text{int}}\{\tau = s + 1 | \mathcal{D}_{\max}\}}{\int_0^{h_0 Q / m_Q} dy \chi(y)}. \quad (3.30)$$

In Fig. 6, we report the distribution of $\tau_{Q,N_s}^{\text{frac}}$ for $s = 5$, $(\alpha, n_{\min}) = (2.45, 3)$ and $Q = 8, \dots, 256$. We observe that it is centered around $t = s$. We also notice that the right tail collapses quickly to $t = s$ as Q increases, while the left tail accumulates towards $t = s$ at a slower pace. As a consequence, backward terms in eq. (3.29) are more quickly suppressed than forward ones. Simulations of $\tau_{Q,Q m_Q^z}^{\text{frac}}$ at $s = 6$ display a rigid shift of the distribution around $t = 6$ with no evident change in the structure of the tails. In Table 4 we report our estimate of $\mathbb{P}_{\text{int}}\{\tau = t | \mathcal{D}_{\max}\}$, a corresponding estimate of the expectation $\mathcal{E}[x | Q, N_s; \mathcal{D}_{\max}]$

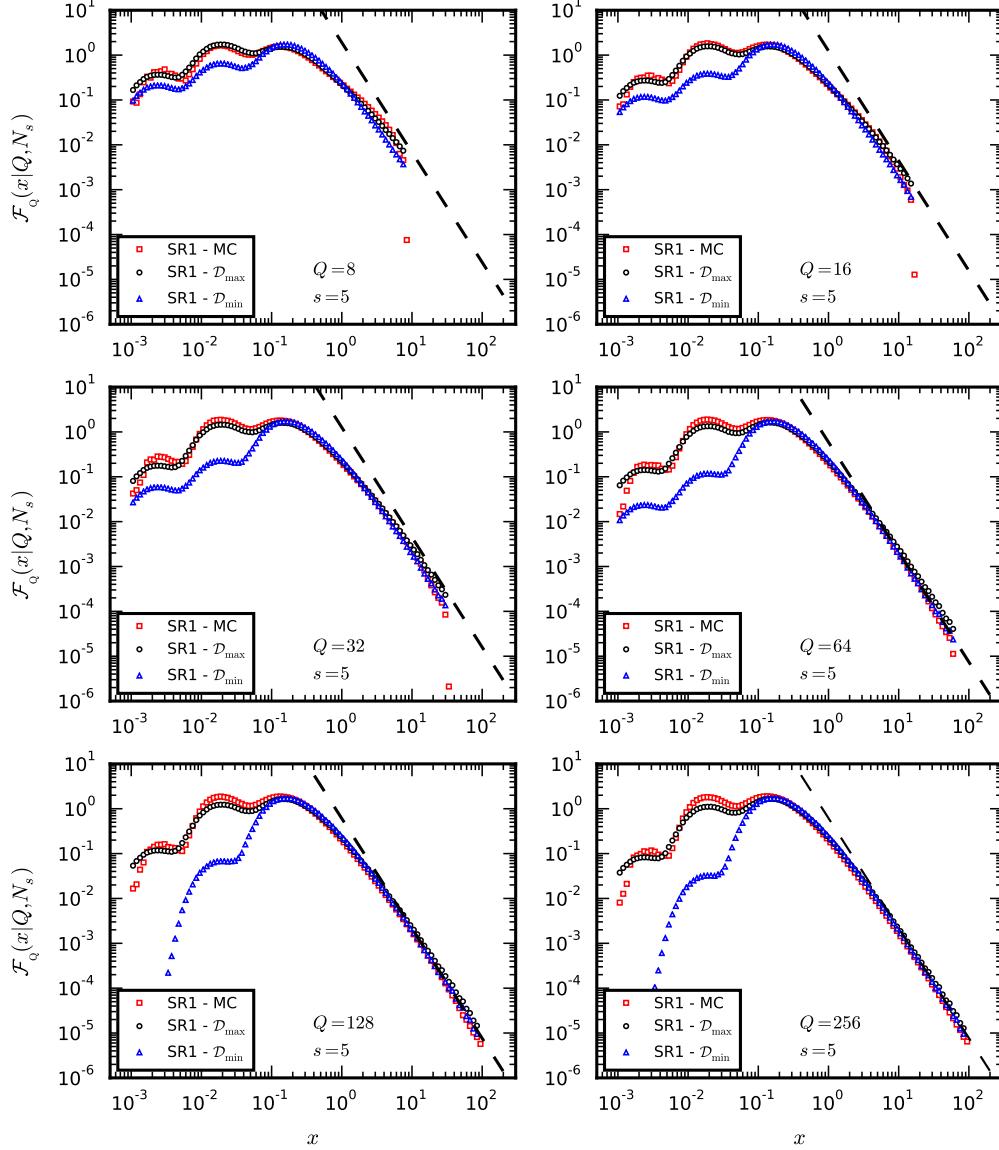


Figure 7 – Monte Carlo simulation (MC) of $\mathcal{F}_Q(x|Q, N_s)$ with stopping rule SR1 vs. our theoretical estimates of the distribution in \mathcal{D}_{\min} and \mathcal{D}_{\max} for $s = 5$, $(\alpha, n_{\min}) = (2.45, 3)$, and $Q = 8, \dots, 256$. The dashed lines represent the asymptotic limit, eq. (3.30).

and the correction coefficient $\mathcal{C}(Q)$ for the same choice of parameters. A comparison with Table 3 shows that $\mathbb{P}_{\text{int}}\{\tau = s | \mathcal{D}_{\max}\} < \mathbb{P}\{\tau = s | \mathcal{D}_{\min}\}$ whereas $\mathbb{P}_{\text{int}}\{\tau = t | \mathcal{D}_{\max}\} > \mathbb{P}\{\tau = t | \mathcal{D}_{\min}\}$ for $t \neq s$. This is reasonable since the system is much less constrained in \mathcal{D}_{\max} than in \mathcal{D}_{\min} . Notably, $\mathcal{E}[x | Q, N_s; \mathcal{D}_{\max}] \rightarrow 1$ from below as $Q \rightarrow \infty$, whereas $\mathcal{E}[x | Q, N_s; \mathcal{D}_{\min}] \rightarrow 1$ from above.

In Fig. 7 we compare Monte Carlo simulations of $\mathcal{F}_Q(x|Q, N_s)$ under SR1 with our theoretical estimates. For all values of Q the numerical distributions are characterized by three bumps along the left tail, slowly decaying as $Q \rightarrow \infty$. These bumps result clearly from the convolution of $\chi(x)$, $\chi(m_Q x)$ and $\chi(m_Q^2 x)$. As can be seen, the bumps are separated by factors of m_Q . This confirms qualitatively the correctness of our analysis. The theoretical distribution in \mathcal{D}_{\max} is in much better agreement with its numerical counterpart than the theoretical distribution in \mathcal{D}_{\min} , specially as $Q \rightarrow \infty$. The dashed lines represent the asymptotic limit of the distribution in \mathcal{D}_{\max} as predicted by eq. (3.30), with the correction coefficient $\mathcal{C}(Q)$ as reported in Table 4.

3.3 $F_Q(x|Q, N)$ with stopping rule SR2

A second case of practical interest is represented by the algorithm with stopping rule SR2. Under this prescription we run eq. (2.4) up to the earliest level t_{exit} at which the overall number of vertices on all trees exceeds N and we arrest the procedure once *all* vertices on the $(t_{\text{exit}} - 1)$ th level have generated their offspring on the t_{exit} th level. In general, we end up with an overall number of vertices $M \geq N$. The *minimal ensemble* \mathcal{D}_{\min} keeps playing an essential rôle in the analytic description of the distribution of the excess-of-votes variable under SR2, yet the adoption of this stopping rule makes the algorithm much less selective than discussed in sects. 3.1 and 3.2. Indeed, the space of all possible vote configurations is now given by

$$\Sigma_+(Q, N) = \left\{ (v_1, \dots, v_Q) \in \mathbb{N}^Q : \sum_{k=1}^Q v_k \geq N \right\} = \bigcup_{M=N}^{\infty} \Sigma(Q, M). \quad (3.31)$$

For convenience we also introduce the complement of $\Sigma_+(Q, N)$ in \mathbb{N}^Q , namely

$$\Sigma_-(Q, N) = \left\{ (v_1, \dots, v_Q) \in \mathbb{N}^Q : \sum_{k=1}^Q v_k < N \right\} = \bigcup_{M=1}^{N-1} \Sigma(Q, M). \quad (3.32)$$

For all $T \in \mathfrak{T}$ there exists $t_+ \geq 1$ such that $(V_{t_+}(T^{(1)}), \dots, V_{t_+}(T^{(Q)})) \in \Sigma_+(Q, N)$. In other words, we have $T \in \bigcup_{M=N}^{\infty} \mathcal{D}_{t_+}(Q, M)$ for some t_+ . Similarly, for all $T \in \mathfrak{T}$ there exists $t_- \geq 0$ such that $(V_{t_-}(T^{(1)}), \dots, V_{t_-}(T^{(Q)})) \in \Sigma_-(Q, N)$, provided $N \geq Q$. In other words, we have $T \in \bigcup_{M=1}^{N-1} \mathcal{D}_{t_-}(Q, M)$ for some t_- . It follows that

$$\bigcup_{t=0}^{\infty} \bigcup_{M=N}^{\infty} \mathcal{D}_t(Q, M) = \bigcup_{M=N}^{\infty} \mathcal{D}_{\min}(Q, M) = \mathfrak{T} \quad \text{and} \quad \bigcup_{t=0}^{\infty} \bigcup_{M=1}^{N-1} \mathcal{D}_t(Q, M) = \bigcup_{M=1}^{N-1} \mathcal{D}_{\min}(Q, M) = \mathfrak{T}, \quad (3.33)$$

independently of N , provided $N \geq Q$. Since we request that the algorithm should stop when the overall number of vertices exceeds N , the stopping time is correctly described by the variable $\tau_{Q,N}$ introduced in eq. (3.6). Nevertheless, different values of $M \geq N$ are realized with different probabilities, hence we must take into account the way $\sum_{i=1}^Q V_{t_{\text{exit}}}^{(i)} = M$ distributes under the condition that $\tau_{Q,N} = t_{\text{exit}}$. Specifically, given (Q, N) and $x = \ell Q/N$ with $\ell = 1, 2, \dots, N$, the discrete probability that the excess-of-votes variable yields x is given by

$$\begin{cases} F_Q(x|Q, N) = \sum_{t=0}^{\infty} F_Q(x|Q, N, t) \mathbb{P}\{\tau_{Q,N} = t\}, \\ F_Q(x|Q, N, t) = \sum_{M=N}^{\infty} \sum_{v \in \Sigma(Q, M)} \delta_{Mx/Q, v_i} \Pi_t(v|Q, M) \mathbb{P}\left\{ \sum_{j=1}^Q V_t^{(j)} = M \middle| \tau_{Q,N} = t \right\}. \end{cases} \quad (3.34)$$

Since in general Mx/Q is not an integer, the above expression has to be interpreted as a linear interpolation between the two distributions obtained by replacing $\delta_{Mx/Q, v_i} \rightarrow \delta_{\lfloor Mx/Q \rfloor, v_i}$ and $\delta_{Mx/Q, v_i} \rightarrow \delta_{\lceil Mx/Q \rceil, v_i}$. The structure of eq. (3.34) suggests that we discuss first $\mathbb{P}\{\tau_{Q,N} = t\}$ and then $F_Q(x|Q, N, t)$. From our analysis in sects. 3.1 and 3.2 we already know that both quantities have to be eventually evaluated for $N \rightarrow N_s = h_0 Q m_Q^s$.

First of all, the probability law of the stopping time $\mathbb{P}\{\tau_{Q,N} = t\}$ is not restricted here to any proper subset of \mathfrak{T} , as it was instead in eqs. (3.7) and (3.24). From eq. (3.33) we see that it amounts to

$$\begin{aligned} \mathbb{P}\{\tau_{Q,N} = t\} &= \mathbb{P}\left\{ \sum_{i=1}^Q V_{t-1}^{(i)} < N, \sum_{i=1}^Q V_t^{(i)} \geq N \right\} \\ &= \mathbb{P}\left\{ T \in \left(\bigcup_{M=1}^{N-1} \mathcal{D}_{t-1}(Q, M) \right) \cap \left(\bigcup_{M=N}^{\infty} \mathcal{D}_t(Q, M) \right) \right\}. \end{aligned} \quad (3.35)$$

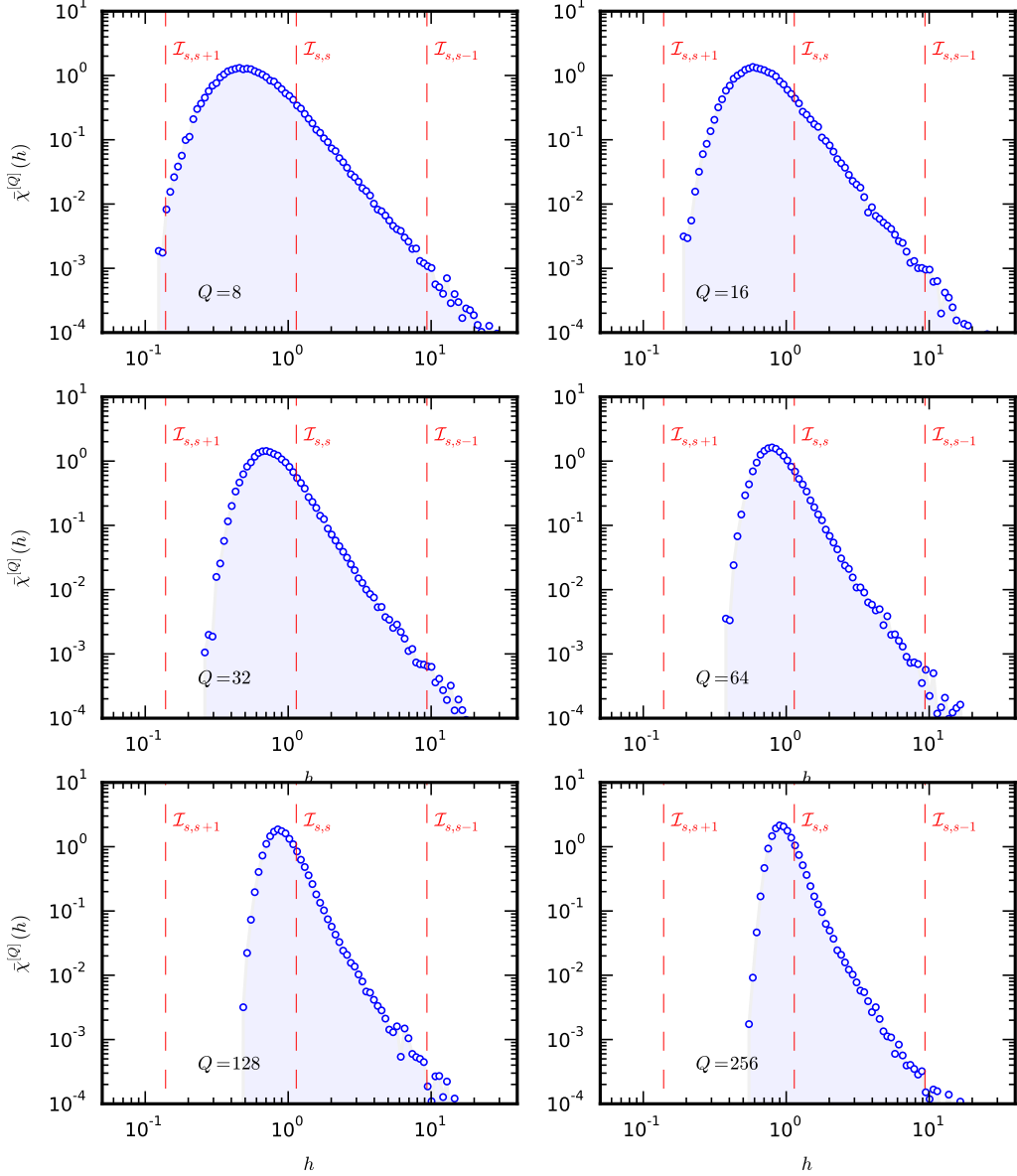


Figure 8 – Distribution of $\bar{H}^{[Q]}$ for $(\alpha, n_{\min}) = (2.45, 3)$ and $Q = 8, \dots, 256$. The dashed lines correspond to $h = h_0 m_Q^{-1}, h_0 m_Q^0, h_0 m_Q^1$. They separate sectors $\mathcal{I}_{s,t}$, $t = s, s \pm 1$.

At first sight, the limit of $\mathbb{P}\{\tau_{Q,N_s} = t\}$ as $s \rightarrow \infty$ is not obvious. To shed light on this, we notice that

$$\begin{aligned} \mathbb{P}\{\tau_{Q,N_s} = t\} &= \mathbb{P}\left\{\sum_{i=1}^Q V_{t-1}^{(i)} < h_0 Q m_Q^s, \sum_{i=1}^Q V_t^{(i)} \geq h_0 Q m_Q^s\right\} \\ &= \mathbb{P}\left\{\bar{H}_{t-1}^{[Q]} < h_0 m_Q^{s-t+1}, \bar{H}_t^{[Q]} \geq h_0 m_Q^{s-t}\right\} \rightarrow \mathbb{P}\left\{h_0 m_Q^{s-t} \leq \bar{H}^{[Q]} < h_0 m_Q^{s-t+1}\right\}, \quad \text{as } s \rightarrow \infty. \end{aligned} \quad (3.36)$$

We thus see that the probability law of the stopping time in the thermodynamic limit is related to the continuous distribution of $\bar{H}^{[Q]}$. More precisely, the domain of the latter splits into contiguous sectors $\mathcal{I}_{s,t} = (h_0 m_Q^{s-t}, h_0 m_Q^{s-t+1}]$ with $t = s, s \pm 1, s \pm 2, \dots$. The mass of $\bar{H}^{[Q]}$ in $\mathcal{I}_{s,t}$ measures the probability of $\tau_{Q,N_s} = t$ as $s \rightarrow \infty$. In Fig. 8 we show the distribution of $\bar{H}^{[Q]}$ obtained from numerical simulations of $\bar{H}_7^{[Q]}$ (which is very close to $\bar{H}^{[Q]}$) for $(\alpha, n_{\min}) = (2.45, 3)$ and $Q = 8, \dots, 256$. The probability mass

$\mathbb{P}\{\bar{H}^{[Q]} \in \mathcal{I}_{s,t}\}$		t			
		$s-2$	$s-1$	s	$s+1$
Q	8	0.00024(6)	0.0064(3)	0.227(2)	0.766(2)
	16	0.00007(5)	0.0042(4)	0.244(3)	0.752(3)
	32	0.00003(2)	0.0033(2)	0.257(2)	0.739(2)
	64	n/a	0.0020(4)	0.271(2)	0.727(2)
	128	n/a	0.0016(1)	0.276(2)	0.722(2)
	256	n/a	0.0012(1)	0.283(2)	0.716(2)
	512	n/a	0.0008(2)	0.285(3)	0.717(4)
	1024	n/a	0.0006(2)	0.293(4)	0.707(4)
	2048	n/a	0.0003(1)	0.298(4)	0.702(4)

Table 5 – Probability of $\bar{H}^{[Q]}$ for $(\alpha, n_{\min}) = (2.45, 3)$ and $Q = 8, \dots, 2048$. Bootstrap uncertainty is quoted in parentheses. N/a entries correspond to measurements for which noise/signal ≥ 1 .

lies essentially in the sectors $\mathcal{I}_{s,s-2}, \mathcal{I}_{s,s-1}, \mathcal{I}_{s,s}, \mathcal{I}_{s,s+1}$. This is analogous to what we observed in sects. 3.1 and 3.2 and shows that the sum over t in eq. (3.34) can be restricted with very good approximation to $t = s-2, s-1, s, s+1$. Secondly, the distribution of $\bar{H}^{[Q]}$ narrows as Q increases. There is a simple explanation for this behaviour. Recall that $\mathbb{E}[\bar{H}^{[Q]}] = h_0$ independently of Q . Consider also that H has a power-law tail $1/h^\alpha$, as eq. (2.14) shows, hence it has infinite variance as long as $\alpha < 3$. Therefore, H is similar to a Paretian variable. The generalized central limit theorem for such variables (see refs. [12, p. 50] and [13, p. 62]) guarantees the existence of a (cumulative) Lévy stable distribution $G(h; \alpha, n_{\min})$ such that

$$\lim_{Q \rightarrow \infty} \mathbb{P} \left\{ \frac{1}{Q^{1/(\alpha-1)}} \left(\sum_{i=1}^Q H^{(i)} - h_0 Q \right) < h \right\} = G(h; \alpha, n_{\min}). \quad (3.37)$$

We have checked numerically the convergence in distribution of $(\sum_{i=1}^Q H^{(i)} - h_0 Q)/Q^{1/(\alpha-1)}$ as $Q \rightarrow \infty$. Since $\alpha - 1 > 1$, an immediate consequence of eq. (3.37) is that

$$\lim_{Q \rightarrow \infty} \bar{\chi}^{[Q]}(h) = \delta(h - h_0). \quad (3.38)$$

In particular, the whole probability mass of $\bar{H}^{[Q]}$ shifts eventually to the sectors $\mathcal{I}_{s,s}$ and $\mathcal{I}_{s,s+1}$ as $Q \rightarrow \infty$. Accordingly, we have

$$\lim_{Q \rightarrow \infty} \lim_{s \rightarrow \infty} \mathbb{P}[\tau_{Q,N_s} = t] = \delta_{t,s} \pi_{s,s} + \delta_{t,s+1} \pi_{s,s+1}, \quad \text{with } \pi_{s,s} + \pi_{s,s+1} = 1. \quad (3.39)$$

It is not easy to determine $\pi_{s,s}$ and $\pi_{s,s+1}$, since the singularity of $\bar{\chi}^{[Q]}(h)$ as $Q \rightarrow \infty$ lies precisely at the common boundary of $\mathcal{I}_{s,s+1}$ and $\mathcal{I}_{s,s}$ (this does not represent a problem for the calculation of $F_Q(x|Q, N)$, as we shall see in the sequel). In Table 5 we quantify the probability of τ_{Q,N_s} in the large-list limit, for $(\alpha, n_{\min}) = (2.45, 3)$ and $Q = 8, \dots, 2048$, via numerical integration of the distributions shown in Fig. 8. At very small Q a large fraction of the probability mass falls in the sector $\mathcal{I}_{s,s+1}$; the rest lies essentially within $\mathcal{I}_{s,s}$, with a residual fraction belonging to $\mathcal{I}_{s,s-1}$ and $\mathcal{I}_{s,s-2}$. Things change smoothly as Q increases in accordance with the above discussion: the probability mass shifts progressively from $\mathcal{I}_{s,s+1}$ to $\mathcal{I}_{s,s}$, with $\mathcal{I}_{s,s-1}$ and $\mathcal{I}_{s,s-2}$ becoming increasingly marginal.

We now go back to the conditional vote distribution $F_Q(x|Q, N, t)$ given t . We have already noticed that it receives contributions from all Q -forests $T \in \bigcup_{M=N}^\infty \mathcal{D}_t(Q, M)$. By performing the same algebra as we did in sect. 3.1, we turn eq. (3.34) into

$$F_Q(x|Q, N, t) = \sum_{M=N}^\infty p_t \left(\frac{Mx}{Q} \right) \left[\frac{\int_{-\pi}^\pi d\lambda e^{-i\lambda \frac{M}{Q}(Q-x)}}{\int_{-\pi}^\pi d\lambda e^{-i\lambda M} \prod_{j=1}^Q \sum_{v_j=0}^M e^{i\lambda v_j} p_t(v_j)} \right] \times \mathbb{P} \left\{ \sum_{j=1}^Q V_t^{(j)} = M \middle| \tau_{Q,N} = t \right\}. \quad (3.40)$$

Provided $x < Q$, we get again highly oscillatory integrals at numerator and denominator of the ratio in square brackets. Therefore, in the large-list limit the above expression converges quickly to

$$F_Q(x|Q, N, t) = \sum_{t=0}^{\infty} \sum_{M=N}^{\infty} \frac{p_t\left(\frac{Mx}{Q}\right)}{\sum_{v_i=0}^M p_t(v_i)} \mathbb{P}\left\{\sum_{j=1}^Q V_t^{(j)} = M \mid \tau_{Q,N} = t\right\}, \quad \text{as } Q, N \rightarrow \infty. \quad (3.41)$$

Now we set $N \rightarrow N_s = h_0 Q m_Q^s$ just as we did in sect. 3.1. We also find it convenient to set $M = N_s + \ell$, with ℓ ranging in principle from zero to infinity. Actually, the probability that $\sum_{j=1}^Q V_t^{(j)} = M$ given that $\tau_{Q,N_s} = t$ is wholly confined in the region $M \leq m_Q N_s$. Indeed, the equation $\sum_{j=1}^Q V_t^{(j)} = M$ is equivalent to $\bar{H}_t^{[Q]} = M/Q m_Q^t$, while the condition $\tau_{Q,N_s} = t$ is equivalent to $h_0 m_Q^{s-t} \leq \bar{H}_t^{[Q]} \leq h_0 m_Q^{s-t+1}$ if t is sufficiently large. It follows that $M \leq h_0 Q m_Q^{s+1} = m_Q N_s$. Therefore, we can restrict the sum over ℓ to the range $\{0, 1, \dots, (m_Q - 1)N_s\}$. From the scaling equation

$$\begin{aligned} p_t(Mx/Q) &= p_t(h_0 x m_Q^s + x\ell/Q) = p_t(m^t [h_0 x m_Q^{s-t} + x\ell/Q m_Q^t]) \\ &\rightarrow (h_0 m_Q^{s-t} + \ell/Q m_Q^t) \chi(h_0 x m_Q^{s-t} + x\ell/Q m_Q^t) dx \quad \text{as } s \rightarrow \infty, \end{aligned} \quad (3.42)$$

we obtain

$$\begin{aligned} \mathcal{F}_Q(x|Q, N_s, t) &= \sum_{\ell=0}^{(m_Q-1)N_s} \left(h_0 m_Q^{s-t} + \frac{\ell}{Q m_Q^t}\right) \frac{\chi\left(h_0 x m_Q^{s-t} + \frac{x\ell}{Q m_Q^t}\right)}{\int_0^{h_0 Q m_Q^{s-t} + \ell/m_Q^t} dy \chi(y)} \\ &\times \mathbb{P}\left\{\sum_{j=1}^Q V_t^{(j)} = N_s + \ell \mid \tau_{Q,N_s} = t\right\}, \quad \text{as } Q, s \rightarrow \infty. \end{aligned} \quad (3.43)$$

The above expression is not in its final form yet, as we are not considering that the discrete sum over ℓ converges to a Riemann integral in the thermodynamic limit. Specifically, from the scaling law

$$\begin{aligned} \mathbb{P}\left\{\sum_{j=1}^Q V_t^{(j)} = N_s + \ell \mid \tau_{Q,N_s} = t\right\} &= \mathbb{P}\left\{\bar{H}_t^{[Q]} = h_0 m_Q^{s-t} + \frac{\ell}{Q m_Q^t} \mid \tau_{Q,N_s} = t\right\} \\ &\rightarrow \bar{\chi}^{[Q]}\left(h_0 m_Q^{s-t} + h \mid \bar{H}^{[Q]} \in \mathcal{I}_{s,t}\right) dh \quad \text{as } s \rightarrow \infty, \end{aligned} \quad (3.44)$$

it follows that

$$\mathcal{F}_Q(x|Q, N_s, t) = \int_{\mathcal{I}_{s,t}} dh \frac{h \chi(hx)}{\int_0^{hQ} dy \chi(y)} \bar{\chi}^{[Q]}\left(h \mid \bar{H}^{[Q]} \in \mathcal{I}_{s,t}\right), \quad \text{as } Q, s \rightarrow \infty. \quad (3.45)$$

The distribution on the r.h.s. of eq. (3.45) is correctly normalized, as can be seen upon integrating both sides over $x \in [0, Q]$. We see from the above equation that the main effect of the stopping rule SR2 is to smooth the distributions $h_0 m_Q^{s-t} \chi(h_0 m_Q^{s-t} x) / \int_0^{h_0 Q m_Q^{s-t}} dy \chi(y)$ contributing to eqs. (3.12) and (3.29) via a convolution with the conditional distribution $\bar{\chi}^{[Q]}(h \mid \bar{H}^{[Q]} \in \mathcal{I}_{s,t})$.

We finally insert eq. (3.45) into eq. (3.34). We notice that since $\bar{\chi}^{[Q]}(h \mid \bar{H}^{[Q]} \in \mathcal{I}_{s,t}) = 0$ if $h \notin \mathcal{I}_{s,t}$, the integral in eq. (3.45) can be extended to $h \in (0, \infty)$ without affecting its value. Accordingly, we obtain our final estimate

$$\begin{aligned} \mathcal{F}_Q(x|Q, N_s) &= \sum_{t=0}^{\infty} \mathcal{F}_Q(x|Q, N_s, t) \mathbb{P}\{\tau_{Q,N_s} = t\} \\ &= \int_0^{\infty} dh \frac{h \chi(hx)}{\int_0^{hQ} dy \chi(y)} \sum_{t=0}^{\infty} \bar{\chi}^{[Q]}\left(h \mid \bar{H}^{[Q]} \in \mathcal{I}_{s,t}\right) \mathbb{P}[\tau_{Q,N_s} = t] \\ &= \int_0^{\infty} dh \frac{h \chi(hx)}{\int_0^{hQ} dy \chi(y)} \bar{\chi}^{[Q]}(h), \quad \text{as } Q, s \rightarrow \infty. \end{aligned} \quad (3.46)$$

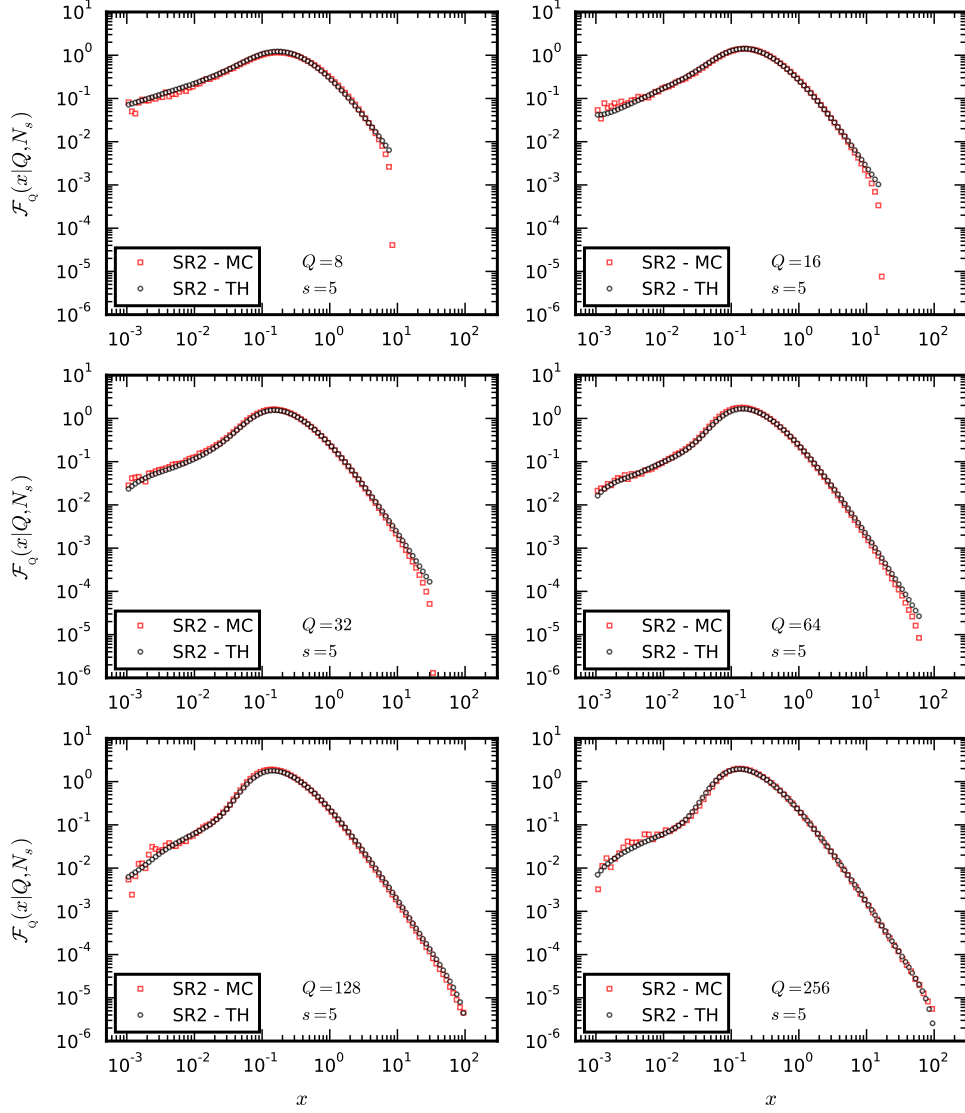


Figure 9 – Monte Carlo simulation (MC) of $\mathcal{F}_Q(x|Q, N_s)$ with stopping rule SR2 vs. theoretical estimate (TH) for $(\alpha, n_{\min}) = (2.45, 3)$, $s = 5$ and $Q = 8, \dots, 256$.

From eq. (3.38) it follows that

$$\lim_{Q \rightarrow \infty} \lim_{s \rightarrow \infty} \mathcal{F}_Q(x|Q, N_s) = h_0 \chi(h_0 x). \quad (3.47)$$

The power-law regime of the limit distribution is reached once more as soon as $x \gtrsim h_0^{-1} \simeq 1$. Eq. (3.47) tells us that the vote distribution in the large-list limit is universal under a change of stopping rule.

Notice that both functions $\chi(x)$ and $\bar{\chi}^{[Q]}(h)$ are known only for a finite number of points, depending on the choice of the bins in the histograms which represent them, see Figs. 5 and 8. Eq. (3.46) can be integrated numerically provided we interpolate $\chi(x)$ and $\bar{\chi}^{[Q]}(h)$ between subsequent observations. We show in Fig. 9 a comparison between the Monte Carlo simulation of $\mathcal{F}_Q(x|Q, N_s)$ and eq. (3.46). The three bumps characterizing the distribution with stopping rule SR1 are now absent due to the convolution with $\bar{\chi}^{[Q]}(h)$. We conclude that only the left tail of the distribution is sensitive to the stopping rule adopted⁶. The agreement between numerical and theoretical results is very good at all scales.

⁶This is analogous to the sensitivity of the left tail to the choice of the model parameters r and k_{\min} , firstly observed by FC in ref. [1].

3.4 $F_Q(x|Q, N)$ with stopping rule SR3

We finally consider the algorithm with stopping rule SR3. According to this prescription, given (Q, s) we stop the generation of new vertices as soon as all variables $(Z_t^{(i)})_{i=1}^Q$ have been realized for $t \leq s$. Under SR3 the stopping time fulfills $\tau_{Q,M} = s$ with certainty, whereas the final number of vertices M on all trees fluctuates freely. In particular, we know that $M \sim O(h_0 Q m_Q^s)$. Therefore, in order to calculate $F_Q(x|Q, s)$ it is sufficient that we properly take the fluctuations of $\sum_{j=1}^Q V_s^{(j)} = M$ into account. Specifically, we let $x_+ = \ell Q / \lfloor h_0 Q m_Q^s \rfloor$ with $\ell = 1, 2, \dots, \lfloor h_0 Q m_Q^s \rfloor$ and $x_- = \ell Q / \lceil h_0 Q m_Q^s \rceil$ with $\ell = 1, 2, \dots, \lceil h_0 Q m_Q^s \rceil$ denote all possible values of the excess-of-votes variable. The discrete probability that the latter yields x is obtained by interpolating the distribution

$$F_Q(x|Q, s) = \sum_{M=0}^{\infty} \sum_{v \in \Sigma(Q, M)} \delta_{Mx/Q, v_i} \Pi_s(v|Q, M) \mathbb{P} \left\{ \sum_{j=1}^Q V_s^{(j)} = M \mid \tau_{Q,M} = s \right\} \quad (3.48)$$

between its values at the nearest points x_{\pm} of x . Since the event $\tau_{Q,M} = s$ implies that $\sum_{j=1}^Q V_s^{(j)} \geq M$, the conditional probability $\mathbb{P} \{ \sum_{j=1}^Q V_s^{(j)} = M \mid \tau_{Q,M} = s \}$ can be safely replaced by its full counterpart, namely $\mathbb{P} \{ \sum_{j=1}^Q V_s^{(j)} = M \}$. Just as in sect. 3.3, the thermodynamic limit is taken by letting $s \rightarrow \infty$ after setting $M = h_0 Q m_Q^s + \ell$. Notice that this time ℓ ranges over the set $\{-h_0 Q m_Q^s, -h_0 Q m_Q^s + 1, \dots, \infty\}$. Repeating the calculation should be at this point trivial and tedious. The reader can easily get convinced that

$$\mathcal{F}_Q(x|Q, s) = \int_0^{\infty} dh \frac{h \chi(hx)}{\int_0^{hQ} dy \chi(y)} \bar{\chi}^{[Q]}(h), \quad \text{as } Q, s \rightarrow \infty. \quad (3.49)$$

We conclude that SR2 and SR3 are equivalent in the large-list limit. Numerical simulations confirm the equivalence with very good accuracy.

4 Exponential scaling in the FC model

The approach we followed to work out the large-list limit of the vote distribution in the quenched model relies ultimately on the exponential scaling of Galton–Watson trees. This led us to perform the thermodynamic limit along the sequence $N_s = h_0 Q m_Q^s$, $s = 1, 2, 3, \dots, \infty$, which corresponds to having the s th level of the Q -forest generated in the algorithm fully realized on average. Given that FC trees are not Galton–Watson, the reader may wonder whether our analysis can be extended to the original FC model. Specifically, since the levels of a FC tree are progressively populated at subsequent times according to eq. (1.1), one could suspect that the overall number of preferences generated at a given time t on a single tree follows a super-exponential behaviour as a function of t . To get confident that this is not the case, it suffices to observe that $Y_n^{(i)}(t) \leq X_n^{(i)}$ for all t and that $X_n^{(i)}$ scales exponentially with n (see Remark 2.1). The exponential growth rate of $V_{\text{FC}}(t)$ can be read off from its expectation value: all we have to do to identify it is to work out the averages $\mathbb{E}[Y_n^{(i)}(t)]$ for $n = 1, 2, \dots, t$ and add them all. Before we embark on this calculation, we recall that in sect. 2 we let m_B denote the average offspring of trees in $\mathcal{P}_B(\alpha, r, k_{\min})$. If, in addition, we let \hat{m}_Q denote the average offspring of trees in $\mathcal{P}_Q(\alpha, k_{\min})$, then we have approximately $m_B = r \hat{m}_Q$. This relation is not exact since it leaves out the effects of fluctuations, yet it is fulfilled with more than acceptable accuracy. We shall use it several times in the next few lines. Incidentally we notice that \hat{m}_Q is just the growth rate of $X_n^{(i)}$, *i.e.* $\mathbb{E}[X_n^{(i)}] = \hat{m}_Q^n$.

In Fig. 10 we show the structure of a generic FC tree at time $t = 5$. The candidate persuades a first group G_1 of agents at time $t = 1$, a second group G_2 at $t = 2$ and so forth. All groups $\{G_k\}_{k=1,2,\dots}$ lie on the 1st level of the tree. Agents belonging to G_1 persuade in turn a first group of agents G_{11} at time $t = 2$,

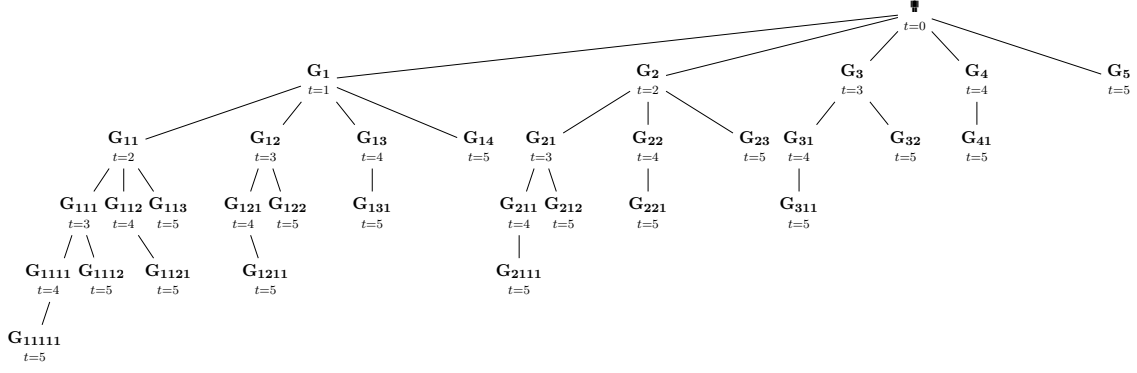


Figure 10 – Structure of a generic FC tree at time $t = 5$.

a second group G_{12} at $t = 3$, *etc.* All groups $\{G_{ij}\}_{i,j=1,2,\dots}$ lie on the 2nd level of the tree. In full generality, agents belonging to one of $\{G_{i_1 i_2 \dots i_n}\}_{i_1, \dots, i_n=1,2,\dots}$ lie on the n th level of the tree; they have been persuaded at time $t = i_1 + \dots + i_n$; given $k \geq 1$, they persuade at time $t + k$ a new group of agents $G_{i_1 i_2 \dots i_n k}$ lying on the $(n + 1)$ th level of the tree.

The average size of G_1 can be calculated at a glance: since the candidate has \hat{m}_Q neighbours on average and each of these is persuaded with probability r , it follows that $\mathbb{E}[|G_1|] = r\hat{m}_Q = m_B$. The average size of G_2 can be easily calculated too: at time $t = 2$ there are on average $(1 - r)\hat{m}_Q$ undecided agents on the 2nd level and each of these is equally persuaded with probability r , hence $\mathbb{E}[|G_2|] = m_B(1 - r)$. By the same argument we can prove that $\mathbb{E}[|G_3|] = m_B(1 - r)^2$, *etc.* In full generality we have $\mathbb{E}[|G_k|] = m_B(1 - r)^{k-1}$ for all groups lying on the 1st level. We can similarly calculate the average size of G_{k1} with $k = 1, 2, \dots$. To this aim we observe that each agent in G_k has \hat{m}_Q neighbours on average and, once more, each of these is persuaded with probability r . Therefore $\mathbb{E}[|G_{k1}|] = m_B \mathbb{E}[|G_k|] = m_B^2(1 - r)^{k-1}$. As a consequence, we have $\mathbb{E}[|G_{k2}|] = r(\hat{m}_Q |G_k| - |G_{k1}|) = m_B^2(1 - r)^k$. We infer that $\mathbb{E}[|G_{ij}|] = m_B^2(1 - r)^{i+j-2}$ holds in general for all groups lying on the 2nd level. It should be clear by now that a simple diagrammatic rule determines $\mathbb{E}[|G_{i_1 \dots i_n}|]$, namely

Each subindex $k = 1, 2, \dots$ contributes to $\mathbb{E}[|G_{i_1 \dots i_n}|]$ by a power of $m_B(1 - r)^{k-1}$.

It follows that

$$\mathbb{E}[|G_{i_1 \dots i_n}|] = m_B^n (1 - r)^{i_1 + \dots + i_n - n}. \quad (4.1)$$

For later convenience, we list below the average size of all groups shown in Fig. 10:

1st level

$$\begin{aligned} \mathbb{E}[|G_1|] &= m_B, & \mathbb{E}[|G_2|] &= m_B(1 - r), & \mathbb{E}[|G_3|] &= m_B(1 - r)^2, \\ \mathbb{E}[|G_4|] &= m_B(1 - r)^3, & \mathbb{E}[|G_5|] &= m_B(1 - r)^4, \end{aligned} \quad (4.2)$$

2nd level

$$\begin{aligned} \mathbb{E}[|G_{11}|] &= m_B^2, & \mathbb{E}[|G_{12}|] &= m_B^2(1 - r), & \mathbb{E}[|G_{13}|] &= m_B^2(1 - r)^2, & \mathbb{E}[|G_{14}|] &= m_B^2(1 - r)^3, \\ \mathbb{E}[|G_{21}|] &= m_B^2(1 - r), & \mathbb{E}[|G_{22}|] &= m_B^2(1 - r)^2, & \mathbb{E}[|G_{23}|] &= m_B^2(1 - r)^3, \\ \mathbb{E}[|G_{31}|] &= m_B^2(1 - r)^2, & \mathbb{E}[|G_{32}|] &= m_B^2(1 - r)^3, & \mathbb{E}[|G_{41}|] &= m_B^2(1 - r)^3, \end{aligned} \quad (4.3)$$

3rd level

$$\begin{aligned}
\mathbb{E}[|G_{111}|] &= m_B^3, \quad \mathbb{E}[|G_{112}|] = m_B^3(1-r), \quad \mathbb{E}[|G_{113}|] = m_B^3(1-r)^2, \\
\mathbb{E}[|G_{121}|] &= m_B^3(1-r), \quad \mathbb{E}[|G_{122}|] = m_B^3(1-r)^2, \quad \mathbb{E}[|G_{131}|] = m_B^3(1-r)^2, \\
\mathbb{E}[|G_{211}|] &= m_B^3(1-r), \quad \mathbb{E}[|G_{212}|] = m_B^3(1-r)^2, \quad \mathbb{E}[|G_{221}|] = m_B^3(1-r)^2, \\
\mathbb{E}[|G_{311}|] &= m_B^3(1-r)^2,
\end{aligned} \tag{4.4}$$

4th level

$$\begin{aligned}
\mathbb{E}[|G_{1111}|] &= m_B^4, \quad \mathbb{E}[|G_{1112}|] = m_B^4(1-r), \quad \mathbb{E}[|G_{1121}|] = m_B^4(1-r), \\
\mathbb{E}[|G_{1211}|] &= m_B^4(1-r), \quad \mathbb{E}[|G_{2111}|] = m_B^4(1-r),
\end{aligned} \tag{4.5}$$

5th level

$$\mathbb{E}[|G_{11111}|] = m_B^5, \tag{4.6}$$

We can use the above expectations to calculate $\mathbb{E}[Y_n^{(i)}(t)]$ for $t = 1, \dots, 5$ and $n \leq t$. Then we infer $\mathbb{E}[Y_n^{(i)}(t)]$ for generic t and n (we leave to the reader the exercise of proving by induction the general formulae given below). We first examine the t th level. We have

$$\begin{aligned}
\mathbb{E}[Y_1^{(i)}(1)] &= \mathbb{E}[|G_1|] = m_B, \\
\mathbb{E}[Y_2^{(i)}(2)] &= \mathbb{E}[|G_{11}|] = m_B^2, \\
\mathbb{E}[Y_3^{(i)}(3)] &= \mathbb{E}[|G_{111}|] = m_B^3, \\
\mathbb{E}[Y_4^{(i)}(4)] &= \mathbb{E}[|G_{1111}|] = m_B^4, \\
\mathbb{E}[Y_5^{(i)}(5)] &= \mathbb{E}[|G_{11111}|] = m_B^5, \\
&\vdots \\
\mathbb{E}[Y_t^{(i)}(t)] &= m_B^t.
\end{aligned} \tag{4.7}$$

We thus see that the t th level of trees in $\mathcal{P}_{\text{FC},t}(\alpha, r, k_{\min})$ scales just like trees in $\mathcal{P}_B(\alpha, r, k_{\min})$. Things get more interesting on the $(t-1)$ th level. In this case we have

$$\begin{aligned}
\mathbb{E}[Y_1^{(i)}(2)] &= \mathbb{E}[|G_1| + |G_2|] = m_B[1 + (1-r)], \\
\mathbb{E}[Y_2^{(i)}(3)] &= \mathbb{E}[|G_{11}| + |G_{12}| + |G_{21}|] = m_B^2[1 + 2(1-r)], \\
\mathbb{E}[Y_3^{(i)}(4)] &= \mathbb{E}[|G_{111}| + |G_{112}| + |G_{121}| + |G_{211}|] = m_B^3[1 + 3(1-r)], \\
\mathbb{E}[Y_4^{(i)}(5)] &= \mathbb{E}[|G_{1111}|] + \mathbb{E}[|G_{1112}|] + \mathbb{E}[|G_{1121}|] + \mathbb{E}[|G_{1211}|] + \mathbb{E}[|G_{2111}|] = m_B^4[1 + 4(1-r)], \\
&\vdots \\
\mathbb{E}[Y_{t-1}^{(i)}(t)] &= [1 + (1-r)(t-1)] \rightarrow (1-r)t m_B^{t-1}, \quad \text{as } t \rightarrow \infty.
\end{aligned} \tag{4.8}$$

We now see that the exponential growth is broken by power corrections. In particular, for the $(t-1)$ th level the correction is linear in t . The additional factor $[1 + (1-r)(t-1)]$ quantifies the effect of the second interaction the undecided agents have with their parents. More generally, we can show that the correction factor to the exponential scaling of persuaded agents on the $(t-k)$ th level amounts to a polynomial in t

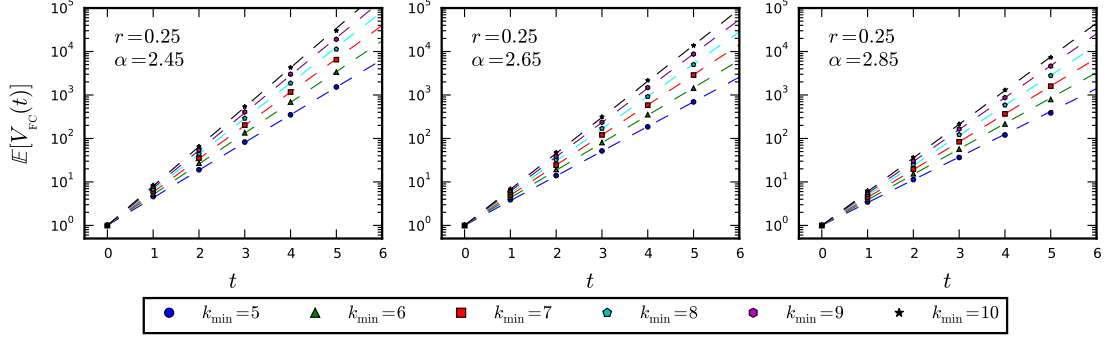


Figure 11 – Scaling of $V_{\text{FC}}(t)$ from numerical simulations. Dashed lines represent eq. (4.10).

of degree k . This quantifies the effect of the k additional interactions the undecided agents have with their parents up to time t . Since we are mainly interested in the scaling of trees in $\mathcal{P}_{\text{FC},t}(\alpha, r, k_{\min})$ as $t \rightarrow \infty$, we keep only the leading term of the polynomial and leave out the subleading ones. Actually, power corrections yield the Taylor expansion of an unknown function of t , which emerges only when we add persuaded agents on all levels. To confirm this, we need to examine at least the $(t-2)$ th level. In this case, we have

$$\begin{aligned}
\mathbb{E}[Y_1^{(i)}(3)] &= \mathbb{E}[|G_1| + |G_2| + |G_3|] = m_{\text{B}}[1 + (1-r) + (1-r)^2], \\
\mathbb{E}[Y_2^{(i)}(4)] &= \mathbb{E}[|G_{11}| + |G_{12}| + |G_{13}| + |G_{21}| + |G_{22}| + |G_{31}|] = m_{\text{B}}^2[1 + 2(1-r) + 3(1-r)^2], \\
\mathbb{E}[Y_3^{(i)}(5)] &= \mathbb{E}[|G_{111}| + |G_{112}| + |G_{113}| + |G_{121}| + |G_{122}| + |G_{131}| + |G_{211}| + |G_{212}| + |G_{221}| + |G_{311}|] \\
&= m_{\text{B}}^3[1 + 3(1-r) + 6(1-r)^2], \\
&\vdots \\
\mathbb{E}[Y_{t-2}^{(i)}(t)] &= m_{\text{B}}^{t-2} \left[1 + (1-r)(t-2) + \frac{1}{2}(1-r)^2(t-1)(t-2) \right] \\
&\rightarrow \frac{1}{2}(1-r)^2 t^2 m_{\text{B}}^{t-2}, \quad \text{as } t \rightarrow \infty.
\end{aligned} \tag{4.9}$$

At this point we can calculate $\mathbb{E}[V_{\text{FC}}(t)]$ as $t \rightarrow \infty$ by just summing eqs. (4.7), (4.8), (4.9), *etc.* In first approximation this yields

$$\begin{aligned}
\mathbb{E}[V_{\text{FC}}(t)] &= \sum_{s=0}^t \mathbb{E}[Y_s^{(i)}(t)] \simeq m_{\text{B}}^t \left\{ 1 + \frac{(1-r)t}{m_{\text{B}}} + \frac{1}{2} \frac{(1-r)^2 t^2}{m_{\text{B}}^2} + \dots \right\} \\
&= \left\{ m_{\text{B}} \exp \left[\frac{(1-r)}{m_{\text{B}}} \right] \right\}^t \equiv m_{\text{FC}}^t, \quad \text{as } t \rightarrow \infty.
\end{aligned} \tag{4.10}$$

Hence, we see that $V_{\text{FC}}(t)$ scales exponentially with growth rate close to $m_{\text{FC}} = m_{\text{B}} \exp\{(1-r)/m_{\text{B}}\}$. In Fig. 11 we show numerical simulations of $\mathbb{E}[V_{\text{FC}}(t)]$ for $r = 0.25$, $\alpha = 2.45, 2.65, 2.85$, $k_{\min} = 5, 6, \dots, 10$ and $t \leq 5$. The dashed lines represent our theoretical estimates. As can be seen, the agreement with simulation data is rather good. In spite of this we must bear in mind that m_{FC} is just an approximation to the true growth rate of $V_{\text{FC}}(t)$. Indeed, as we noticed above, to resum in closed form the power corrections we had to drop all subleading terms on each level. By doing this, we did not consider that each subleading term mixes with the leading term of some upper level, thus producing a small shift of its coefficient proportional to some inverse power of m_{B} . We conclude that m_{FC} represents the correct growth rate up to $\mathcal{O}(1/m_{\text{B}})$. We shall come back to this in a while.

In consideration of eq. (4.10), we find it convenient to introduce the rescaled variable

$$H_{\text{FC}}(t) = \frac{V_{\text{FC}}(t)}{m_{\text{FC}}^t}. \quad (4.11)$$

From the above discussion it follows that $\mathbb{E}[H_{\text{FC}}(t)] = 1 + \mathcal{O}(1/m_{\text{B}})$. As $t \rightarrow \infty$ the sequence $(H_{\text{FC}}(t))_{t \geq 0}$ converges to a finite limit H_{FC} with continuous *p.d.f.* $\chi_{\text{FC}}(h)$. To prove this, it suffices to show that $H_{\text{FC}}(t)$ is a martingale. To this aim we need to evaluate $\mathbb{E}[V_{\text{FC}}(t) | \{G_{i_1 \dots i_n}\}_{i_1 + \dots + i_n \leq t-1}]$, *i.e.* the conditional expectation of $V_{\text{FC}}(t)$ given all groups which have been generated up to time $t-1$. As previously, we perform the calculation for $t = 1, 2, \dots$ and then we extrapolate to generic t . With reference to Fig. 10, we have

$$\begin{aligned} V_{\text{FC}}(1) &= 1 + |G_1|, \\ V_{\text{FC}}(2) &= 1 + |G_1| + |G_2| + |G_{11}|, \\ V_{\text{FC}}(3) &= 1 + |G_1| + |G_2| + |G_3| + |G_{11}| + |G_{12}| + |G_{21}| + |G_{111}|, \\ &\vdots \end{aligned} \quad (4.12)$$

At time $t = 2$ the conditional expectation yields

$$\begin{aligned} \mathbb{E}[V_{\text{FC}}(2) | G_1] &= 1 + |G_1| + \mathbb{E}[|G_2| | G_1] + \mathbb{E}[|G_{11}| | G_1] = 1 + |G_1| + r(\hat{m}_{\text{Q}} - |G_1|) + m_{\text{B}}|G_1| \\ &= [m_{\text{B}} + (1 - r)]V_{\text{FC}}(1) + r. \end{aligned} \quad (4.13)$$

At time $t = 3$ it yields

$$\begin{aligned} \mathbb{E}[V_{\text{FC}}(3) | G_1, G_2, G_{11}] &= 1 + |G_1| + |G_2| + |G_{11}| \\ &+ \mathbb{E}[|G_3| | G_1, G_2] + \mathbb{E}[|G_{12}| | G_1, G_{11}] + \mathbb{E}[|G_{21}| | G_2] + \mathbb{E}[|G_{111}| | G_{11}] \\ &= 1 + |G_1| + |G_2| + |G_{11}| + r(\hat{m}_{\text{Q}} - |G_1| - |G_2|) + r(\hat{m}_{\text{Q}}|G_1| - |G_{11}|) + m_{\text{B}}|G_2| + m_{\text{B}}|G_{11}| \\ &= [m_{\text{B}} + (1 - r)]V_{\text{FC}}(2) + r. \end{aligned} \quad (4.14)$$

At time $t = 4$ we can equally show that $\mathbb{E}[V_{\text{FC}}(4) | \{G_{i_1 \dots i_n}\}_{i_1 + \dots + i_n \leq 3}] = [m_{\text{B}} + (1 - r)]V_{\text{FC}}(3) + r$. Hence, we conclude that the equation

$$\mathbb{E}[V_{\text{FC}}(t) | \{G_{i_1 \dots i_n}\}_{i_1 + \dots + i_n \leq t-1}] = [m_{\text{B}} + (1 - r)]V_{\text{FC}}(t-1) + r \quad (4.15)$$

holds for generic t . Notice that $m_{\text{B}} + (1 - r) = m_{\text{FC}} + \mathcal{O}(1/m_{\text{B}})$. Since the additive term r on the r.h.s. of eq. (4.15) represents an infinitesimal correction as $t \rightarrow \infty$, we conclude that $H_{\text{FC}}(t)$ is asymptotically a martingale. In Fig. 12 we compare $\chi_{\text{FC}}(h)$ with $\chi(h)$. The two *p.d.f.*'s look very similar for $h \geq 1$. Their tails have the same power-law exponent and essentially the same scale factor. As we know, the latter is sufficiently well approximated by $h_0^{\alpha-1}/\zeta(\alpha-1, n_{\text{min}})$. In practice, they differ only for $h \leq 1$, where the effects of the activism of the agents and the noise of the agent-agent interactions become important. To conclude, we observe that the analysis presented in sect. 3 can now be easily adapted to the FC model, provided we replace $m_{\text{Q}} \rightarrow m_{\text{FC}}$, $\chi(h) \rightarrow \chi_{\text{FC}}(h)$ and $h_0 \rightarrow 1$.

5 One-step transition probabilities

We now go back to the distribution of W . In sect. 2 we wrote eq. (2.10) without explaining where it comes from. In principle $\phi(w)$ can be derived from $\mathbb{P}\{Z_t = z\}$ by setting $z = wm_{\text{Q}}^t$ and by using the scaling law $\mathbb{P}\{Z_t = wm_{\text{Q}}^t\} = \mathbb{P}\{W_t = w\} \rightarrow \phi(w)dw$ as $t \rightarrow \infty$. Since $\mathbb{E}[Z_t] = m_{\text{Q}}^t$, the tail the distribution is observed

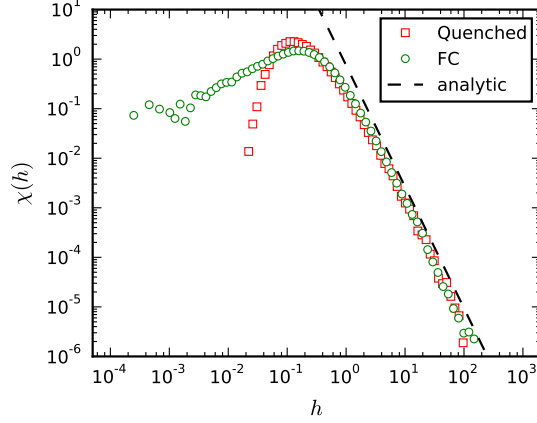


Figure 12 – $\chi(h)$ for $(\alpha, n_{\min}) = (2.45, 3)$ vs. $\chi_{\text{FC}}(h)$ for $(\alpha, r, k_{\min}) = (2.45, 0.25, 10)$.

for $z \gg m_Q^t$, that is for $w \gg 1$. In order to work out $\phi(w)$ in this limit, we adopt a strategy based on the one-step transition probabilities (OSTP) $\mathbb{P}\{Z_t = z | Z_{t-1}\}$. Although these objects are seldom used for the analysis of branching trees⁷, they can be regarded to all purposes as building blocks for theoretical calculations, in that they propagate complete information from one level to the next one. Specifically, we have

$$\mathbb{P}\{Z_t = z\} = \sum_{z_{t-1}, \dots, z_1} \mathbb{P}\{Z_t = z | Z_{t-1} = z_{t-1}\} \cdot \mathbb{P}\{Z_{t-1} = z_{t-1} | Z_{t-2} = z_{t-2}\} \cdot \dots \cdot \mathbb{P}\{Z_1 = z_1\}, \quad (5.1)$$

$$\mathbb{P}\{V_t = v\} = \sum_{\substack{z_t, \dots, z_1 \\ |z| = v-1}} \mathbb{P}\{Z_t = z_t | Z_{t-1} = z_{t-1}\} \cdot \mathbb{P}\{Z_{t-1} = z_{t-1} | Z_{t-2} = z_{t-2}\} \cdot \dots \cdot \mathbb{P}\{Z_1 = z_1\}. \quad (5.2)$$

Here we are interested in particular in eq. (5.1), which is also equivalent to

$$\mathbb{P}\{Z_t = z\} = \sum_{\tilde{z}} \mathbb{P}\{Z_t = z | Z_{t-1} = \tilde{z}\} \cdot \mathbb{P}\{Z_{t-1} = \tilde{z}\} = \mathbb{E}[\mathbb{P}\{Z_t = z | Z_{t-1}\}]. \quad (5.3)$$

The law of iterated expectations is useful only if we are able to both calculate $\mathbb{P}\{Z_t = z | Z_{t-1} = \tilde{z}\}$ and average it over \tilde{z} . Before we embark on this, we recall that $Z_t = \sum_{j=1}^{Z_{t-1}} \xi_j$ with $\xi_j \sim p_Q(n)$ and $Z_{t-1} \geq (n_{\min})^{t-1}$. Therefore, if $n_{\min} > 1$ and t is not too small, Z_t is the sum of a large number of i.i.d. Mandelbrot variables. Since ξ_j has infinite variance for $\alpha < 3$, from the generalized central limit theorem (see refs. [12, p. 50] and [13, p. 62]) it follows that $\mathbb{P}\{Z_t = z | Z_{t-1} = \tilde{z}\}$ falls in the domain of attraction of a stable law of index $(\alpha - 1)$ as $\tilde{z} \rightarrow \infty$, *i.e.* there exists a Lévy stable distribution $\Sigma(z; \alpha, n_{\min})$ with *p.d.f.* $\sigma(z; \alpha, n_{\min})$ such that

$$\lim_{\tilde{z} \rightarrow \infty} \mathbb{P}\left\{\frac{1}{\tilde{z}^{1/(\alpha-1)}} \left(\sum_{j=1}^{\tilde{z}} \xi_j - m_Q \tilde{z}\right) < z\right\} = \Sigma(z; \alpha, n_{\min}). \quad (5.4)$$

In Fig. 13 (left) we plot the distribution of $(\sum_{j=1}^{\tilde{z}} \xi_j - m_Q \tilde{z})/\tilde{z}^{1/(\alpha-1)}$ for $\alpha = 2.45$, $n_{\min} = 3$ and $\tilde{z} = 10^1, \dots, 10^4$. The plot shows that the convergence to the limit distribution $\sigma(z; \alpha, n_{\min})$ is very fast. In view of this, we conclude that $\mathbb{P}\{Z_t = z | Z_{t-1} = \tilde{z}\}$ has a power-law tail as a function of z even for moderate values of \tilde{z} . Although we can calculate the OSTP, we are not able to work out its expectation in closed form. In first approximation, we can get the tail of the full probability $\mathbb{P}\{Z_t = z\}$ by expanding $\mathbb{P}\{Z_t = z | Z_{t-1} = \tilde{z}\}$ asymptotically in inverse powers of z as $z \rightarrow \infty$ and by then averaging over \tilde{z} just the leading term of the expansion. We anticipate that the resulting estimate is not uniformly good in α . In

⁷Most known results are indeed based on the use of probability generating functions.

particular, our estimate is mathematically consistent with the scaling law of $\mathbb{P}\{Z_t = z\}$ only for $\alpha \rightarrow 2$. In this limit eq. (2.10) holds true. We shall see, however, that the error we make for $\alpha > 2$ is not exceedingly large.

In order to perform the calculation of $\mathbb{P}\{Z_t = z | Z_{t-1} = \tilde{z}\}$, we follow ref. [14], where a similar study is carried out for (continuous) Paretian variables. In fact, the only difference with that paper is that our variables are discrete. We introduce preliminarily the Lerch transcendent

$$\Phi(z, \alpha, n) = \sum_{j=0}^{\infty} \frac{z^j}{(j+n)^\alpha}. \quad (5.5)$$

This special function generalizes the Hurwitz ζ -function, which is indeed obtained for $z = 1$. We notice that $\Phi(z, \alpha, n)$ is analytic for $|z| < 1$ and, if $\text{Re}\{\alpha\} > 1$, also for $|z| = 1$, provided $n \neq 0, -1, -2, \dots$. For the other values of z , $\Phi(z, \alpha, n)$ is defined by analytic continuation. In particular, $\Phi(z, \alpha, n)$ can be differentiated infinitely many times at $z = 0$. It is also useful to recall that $\Phi(z, \alpha, n)$ can be represented as a Taylor series, namely (see ref. [15, p. 29])

$$\Phi(z, \alpha, n) = z^{-n} \left[\Gamma(1-\alpha) [-\log(z)]^{\alpha-1} + \sum_{a=0}^{\infty} \zeta(\alpha-a, n) \frac{\log(z)^a}{a!} \right]. \quad (5.6)$$

This series converges for $|\log(z)| < 2\pi$, $\alpha \neq 1, 2, 3, \dots$ and $n \neq 0, -1, -2, \dots$. The special character of $\Phi(z, \alpha, n)$ is clearly expressed by the presence of a fractional power in the Taylor series. For $\alpha < 3$ the term $[-\log(z)]^{\alpha-1}$ is the ultimate source of the power-law behaviour of $\mathbb{P}\{Z_t = z | Z_{t-1} = \tilde{z}\}$, as we shall see in the next few lines.

Since we have $Z_t = \sum_{j=1}^{Z_{t-1}} \xi_j$, the OSTP can be written as a convolution of $Z_{t-1} = \tilde{z}$ copies of the Mandelbrot distribution, namely

$$\begin{aligned} \mathbb{P}\{Z_t = z | Z_{t-1} = \tilde{z}\} &= \sum_{n_1, \dots, n_{\tilde{z}}=0}^{\infty} \delta_{z, n_1 + \dots + n_{\tilde{z}}} p_Q(n_1) \dots p_Q(n_{\tilde{z}}) \\ &= [\zeta(\alpha, n_{\min})]^{-\tilde{z}} \sum_{n_1, \dots, n_{\tilde{z}}=0}^{\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} d\phi e^{i\phi(n_1 + \dots + n_{\tilde{z}} - z)} \frac{\theta_{n_1, n_{\min}}}{n_1^\alpha} \dots \frac{\theta_{n_{\tilde{z}}, n_{\min}}}{n_{\tilde{z}}^\alpha} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} d\phi e^{-i\phi(z - \tilde{z} n_{\min})} \left[\frac{\Phi(e^{i\phi}, \alpha, n_{\min})}{\zeta(\alpha, n_{\min})} \right]^{\tilde{z}}, \end{aligned} \quad (5.7)$$

where we used the Fourier representation of the Kronecker delta and we let θ denote the discrete Heaviside step function, namely $\theta_{a,b} = 1$ if $a \geq b$ and $\theta_{a,b} = 0$ otherwise. We insert the Taylor expansion of the Lerch transcendent into eq. (5.7). This yields the expression

$$\mathbb{P}\{Z_t = z | Z_{t-1} = \tilde{z}\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\phi e^{-i\phi z} \left[\frac{\Gamma(1-\alpha)}{\zeta(\alpha, n_{\min})} \left(\frac{\phi}{i} \right)^{\alpha-1} + \sum_{a=0}^{\infty} \frac{(-1)^a}{a!} \frac{\zeta(\alpha-a, n_{\min})}{\zeta(\alpha, n_{\min})} \left(\frac{\phi}{i} \right)^a \right]^{\tilde{z}}. \quad (5.8)$$

Now we observe that if $2 < \alpha < 3$ the fractional power $(\phi/i)^{\alpha-1}$ lies between the integer powers (ϕ/i) and $(\phi/i)^2$, while if $\alpha > 3$ it lies between $(\phi/i)^2$ and $(\phi/i)^3$. The first integer power (ϕ/i) is related to the conditional expectation $\mathbb{E}[Z_t | Z_{t-1}]$, as can be seen from its coefficient which is precisely $-m_Q$. To take advantage of this, we insert $1 = \exp\{i\phi m_Q \tilde{z}\} \exp\{-i\phi m_Q \tilde{z}\}$ under the integral sign and then expand the exponential $\exp\{-i\phi m_Q \tilde{z}\} = (\exp\{-i\phi m_Q\})^{\tilde{z}}$ in Taylor series. Accordingly, we obtain

$$\mathbb{P}\{Z_t = z | Z_{t-1} = \tilde{z}\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\phi e^{-i\phi(z - m_Q \tilde{z})} \left[1 + \frac{\Gamma(1-\alpha)}{\zeta(\alpha, n_{\min})} \left(\frac{\phi}{i} \right)^{\alpha-1} + O(\phi^2) \right]^{\tilde{z}}, \quad (5.9)$$

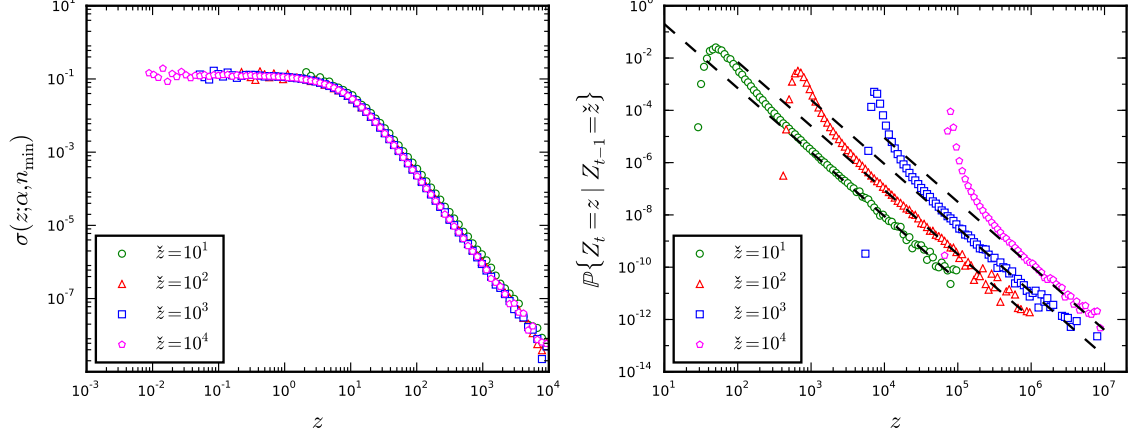


Figure 13 – (Left) *p.d.f.* $\sigma(z; \alpha, n_{\min})$ for $\alpha = 2.45$ and $n_{\min} = 3$ as obtained from numerical simulations of $[\sum_{j=1}^{\tilde{z}} \xi_j - m_Q \tilde{z}] / \tilde{z}^{1/(\alpha-1)}$ with $\tilde{z} = 10^1, 10^2, 10^3, 10^4$; (Right) OSTP for the same choice of parameters. The dashed lines represent eq. (5.19).

for $2 < \alpha < 3$, and

$$\mathbb{P}\{Z_t = z \mid Z_{t-1} = \tilde{z}\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\phi e^{-i\phi(z - m_Q \tilde{z})} \left[1 + \frac{1}{2} \frac{\zeta(\alpha - 2, n_{\min})}{\zeta(\alpha, n_{\min})} \left(\frac{\phi}{i} \right)^2 + o(\phi^2) \right]^{\tilde{z}}, \quad (5.10)$$

for $\alpha > 3$. We are interested in how the OSTP looks as $\tilde{z} \rightarrow \infty$. For this, we rescale the integration variable according to $\phi \rightarrow \tilde{z}^{1/(\alpha-1)} \phi$ if $2 < \alpha < 3$ and according to $\phi \rightarrow \tilde{z}^{1/2} \phi$ if $\alpha > 3$, then we use the notable limit $\lim_{x \rightarrow \infty} (1 + \lambda/x)^x = \exp\{\lambda\}$ and finally we scale back the integration variable. This yields

$$\begin{aligned} \mathbb{P}\{Z_t = z \mid Z_{t-1} = \tilde{z}\} &= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\phi \exp \left\{ -i\phi(z - m_Q \tilde{z}) + \frac{\Gamma(1 - \alpha)}{\zeta(\alpha - 1, n_{\min})} m_Q \tilde{z} \left(\frac{\phi}{i} \right)^{\alpha-1} \right\} \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\phi e^{-i\phi z} \\ &\quad \cdot \exp \left\{ i\phi m_Q \tilde{z} + \frac{\Gamma(1 - \alpha)}{\zeta(\alpha - 1, n_{\min})} m_Q \tilde{z} \cos \left(\frac{\pi}{2} (\alpha - 1) \right) |\phi|^{\alpha-1} \left[1 - i \operatorname{sign}(\phi) \tan \left(\frac{\pi}{2} (\alpha - 1) \right) \right] \right\}, \end{aligned} \quad (5.11)$$

for $2 < \alpha < 3$ and

$$\mathbb{P}\{Z_t = z \mid Z_{t-1} = \tilde{z}\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\phi e^{-i\phi z} \exp \left\{ i\phi m_Q \tilde{z} - \frac{1}{2} \frac{\zeta(\alpha - 2, n_{\min})}{\zeta(\alpha - 1, n_{\min})} m_Q \tilde{z} \phi^2 \right\}, \quad (5.12)$$

for $\alpha > 3$. A comment is in order concerning the integration limits. From eq. (5.7) we know that $\mathbb{P}\{Z_t = z \mid Z_{t-1} = \tilde{z}\} = 0$ unless $z \geq n_{\min} \tilde{z}$, therefore the probability mass of the OSTP shifts progressively towards $z \rightarrow \infty$ as $\tilde{z} \rightarrow \infty$. Moreover, since $Z_t \in \mathbb{N}$ its minimum variation amounts to $\Delta Z_t = 1$ independently of Z_t . It follows that $\Delta Z_t / Z_t = 1/z \rightarrow 0$ as $\tilde{z} \rightarrow \infty$, hence $\mathbb{P}\{Z_t = z \mid Z_{t-1} = \tilde{z}\}$ becomes to all practical purposes a continuous distribution in z . This explains why the integral transform in eqs. (5.11)–(5.12) is over $\phi \in (-\infty, \infty)$.

If we compare the second exponential in eq. (5.11) with the general expression of the characteristic function of a stable distribution of index $(\alpha - 1)$ given in ref. [16, p. 164], namely

$$\Psi(\phi) = \exp \left\{ i\gamma\phi - c|\phi|^{\alpha-1} \left[1 + i\beta \operatorname{sign}(\phi) \tan \left(\frac{\pi}{2} (\alpha - 1) \right) \right] \right\}, \quad (5.13)$$

we see that

$$\gamma = m_Q \tilde{z}, \quad c = -\frac{\Gamma(1 - \alpha)}{\zeta(\alpha - 1, n_{\min})} m_Q \tilde{z} \cos \left(\frac{\pi}{2} (\alpha - 1) \right), \quad \beta = -1. \quad (5.14)$$

In particular, γ is the location parameter of the distribution (its value here is in agreement with $\mathbb{E}[Z_t | Z_{t-1} = m_Q Z_{t-1}] = m_Q Z_{t-1}$), c is the scale parameter (it is proportional to γ in our case) and β represents the skewness of the distribution (in general β fulfills $|\beta| \leq 1$, hence it takes here the maximum negative value). The second exponential in eq. (5.12) is the characteristic function of a normal variable with mean $\mu = m_Q \tilde{z}$ and variance $\sigma^2 = m_Q \tilde{z} \zeta(\alpha - 2, n_{\min}) / \zeta(\alpha - 1, n_{\min})$. Since in the quenched model we are always interested in $2 < \alpha < 3$, we focus on eq. (5.11) and forget about eq. (5.12) in the following.

5.1 Tail of the distribution

A general formula describing the asymptotic behaviour of the *p.d.f.* of a Lévy stable distribution is given in ref. [17, ch. 1]. For the sake of completeness (and to be sure that the reader interprets correctly the formula), we reproduce it here. We start from eq. (5.11), that we recast in the form

$$\mathbb{P}\{Z_t = z | Z_{t-1} = \tilde{z}\} = \frac{1}{\pi} \operatorname{Re} \left\{ \int_0^\pi d\phi e^{i\phi(z - m_Q \tilde{z})} e^{-\rho \phi^{\alpha-1}} \right\}, \quad (5.15)$$

with $\rho = -\frac{\Gamma(1-\alpha)}{\zeta(\alpha-1, n_{\min})} \exp\{-i\frac{\pi}{2}(\alpha-1)\} m_Q \tilde{z}$. We expand the exponential $\exp\{-\rho \phi^{\alpha-1}\}$ in Taylor series. Then, we integrate the series term by term. In this way, we obtain

$$\begin{aligned} \mathbb{P}\{Z_t = z | Z_{t-1} = \tilde{z}\} &= \frac{1}{\pi} \operatorname{Re} \left\{ \sum_{\ell=0}^{\infty} \frac{(-1)^\ell \rho^\ell}{\ell!} \int_0^\pi d\phi \phi^{\ell(\alpha-1)} e^{-i(z - m_Q \tilde{z})\phi} \right\} \\ &= \frac{1}{\pi} \operatorname{Re} \left\{ \sum_{\ell=0}^{\infty} \frac{(-1)^\ell \rho^\ell}{\ell!} \frac{\Gamma(1 + (\alpha-1)\ell)}{[i(z - m_Q \tilde{z})]^{1+\ell(\alpha-1)}} \right\}. \end{aligned} \quad (5.16)$$

Apart from the first term of the series, which is a Dirac delta $\delta(z - m_Q \tilde{z})$, all the subsequent terms are inverse powers of $(z - m_Q \tilde{z})$. Therefore, at leading order the expansion reads

$$\mathbb{P}\{Z_t = z | Z_{t-1} = \tilde{z}\} = \delta(z - m_Q \tilde{z}) - \frac{1}{\pi} \frac{\Gamma(\alpha) \operatorname{Re}\{\rho e^{-i\alpha\pi/2}\}}{(z - m_Q \tilde{z})^\alpha} + \mathcal{O}\left(\frac{1}{(z - m_Q \tilde{z})^{2\alpha-1}}\right). \quad (5.17)$$

For $z \gg m_Q \tilde{z}$ we can drop both the Dirac delta and the subleading terms of the expansion. Moreover, from the Euler reflection formula $\Gamma(\alpha)\Gamma(1-\alpha) = \pi/\sin(\pi\alpha)$, it follows that

$$\frac{\Gamma(\alpha) \operatorname{Re}\{\rho e^{-i\alpha\pi/2}\}}{\pi} = -\frac{1}{\pi} \frac{\Gamma(\alpha)\Gamma(1-\alpha)}{\zeta(\alpha-1, n_{\min})} m_Q \tilde{z} \operatorname{Re}\{e^{-i\pi(\alpha-1/2)}\} = -\frac{m_Q \tilde{z}}{\zeta(\alpha-1, n_{\min})}. \quad (5.18)$$

Therefore, we end up with the asymptotic estimate

$$\mathbb{P}\{Z_t = z | Z_{t-1} = \tilde{z}\} = \frac{1}{\zeta(\alpha-1, n_{\min})} \frac{m_Q \tilde{z}}{z^\alpha}, \quad \text{as } z \rightarrow \infty. \quad (5.19)$$

In Fig. 13 (right) we show the OSTP for $\alpha = 2.45$, $n_{\min} = 3$ and $\tilde{z} = 10^1, \dots, 10^4$. The dashed lines on the plot represent the power-law tail as predicted by eq. (5.19). We see that the theoretical estimates are in perfect agreement with the simulations.

Now, averaging eq. (5.19) over $Z_{t-1} = \tilde{z}$ yields

$$\mathbb{P}\{Z_t = z\} = \frac{1}{\zeta(\alpha-1, n_{\min})} \frac{m_Q^t}{z^\alpha}, \quad \text{as } z \rightarrow \infty. \quad (5.20)$$

As anticipated, this formula is inconsistent with the scaling law of $\mathbb{P}\{Z_t = z\}$ unless $\alpha \rightarrow 2$. Indeed, by setting $z = w m_Q^t$, we get

$$\mathbb{P}\{W_t = w\} = \frac{1}{m_Q^{(\alpha-1)t}} \frac{1}{\zeta(\alpha-1, n_{\min})} \frac{1}{w^\alpha}, \quad \text{as } w \rightarrow \infty. \quad (5.21)$$

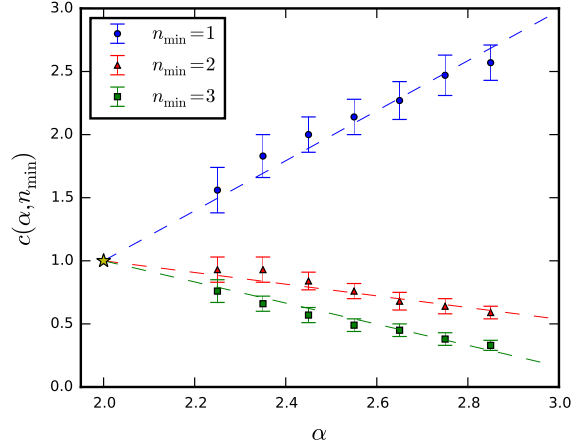


Figure 14 – Coefficient $c(\alpha, n_{\min})$ (see eq. (5.22)) as obtained from numerical simulations of the distribution of W_6 for $\alpha = 2.25, \dots, 2.85$ and $n_{\min} = 1, 2, 3$.

Since the minimum variation of W_t amounts to $\Delta W_t = 1/m_Q^t \rightarrow dw$ as $t \rightarrow \infty$, we conclude that eq. (5.21) scales anomalously with $dw^{\alpha-1}$. This is not totally surprising in consideration that *i*) we obtained our estimate from a representation of the OSTP which formally holds only in the limit $\tilde{z} \rightarrow \infty$ and *ii*) we dropped additional terms in the OSTP in the limit $z \rightarrow \infty$. Notice, however, that the scaling anomaly disappears as $\alpha \rightarrow 2$. In this limit eq. (5.21) scales correctly with dw . It is only for $\alpha \gtrsim 2$ that we expect $\phi(w)$ to be accurately described by eq. (2.10). If we make the ansatz

$$\phi(w) = \frac{c(\alpha, n_{\min})}{\zeta(\alpha - 1, n_{\min})} \frac{1}{w^\alpha}, \quad \text{as } w \rightarrow \infty, \quad (5.22)$$

we can measure the constant $c(\alpha, n_{\min})$ by means of numerical simulations. In Fig. 14 we report our determinations of $c(\alpha, n_{\min})$ for $\alpha = 2.25, \dots, 2.85$ and $n_{\min} = 1, 2, 3$ from simulations of W_6 . The plot confirms that $c(\alpha, n_{\min}) \rightarrow 1$ as $\alpha \rightarrow 2$. Moreover, it shows that $1/2 \lesssim c(\alpha, n_{\min}) \lesssim 2$ for $\alpha \lesssim 2.5$ and $n_{\min} \leq 3$. For these values of the model parameters eq. (2.10) provides a reasonably good approximation.

To conclude, we stress once more that the above derivation relies on eq. (5.11), which holds formally in the limit $\tilde{z} \rightarrow \infty$. The reader may wonder how difficult it would be to calculate the OSTP for generic \tilde{z} . To answer this question, we derive in App. B an alternative representation of $\mathbb{P}\{Z_t = z \mid Z_{t-1} = \tilde{z}\}$, based on multiple polylogarithms and shuffle products, as a series of inverse powers $z^{-\alpha_1}$ with $\alpha_1 = \alpha, \alpha + 1, \dots$. This representation is valid for integer α and makes no assumptions on \tilde{z} .

6 Discussion and outlook

The word-of-mouth model for proportional elections, proposed by Fortunato and Castellano in ref. [1], reproduces with great accuracy the scaling distribution $\mathcal{F}_{\text{EMP}}(x)$, universally observed in elections held in different countries and years. As such, the model represents a significant breakthrough in the field of opinion dynamics, where a qualitative agreement between empirical data and theoretical models is often the best result one can achieve. In spite of this, the analytic structure of the model had never been studied so far, neither by its authors nor by other scholars, and the only available information was based on computer simulations. In the present paper we made a first step to fill this gap.

It was known, in particular, that the conditional distribution $F_{\text{FC}}(x|Q, N)$ predicted by the model develops a power-law right tail as $Q \rightarrow \infty$. The amount of empirical observations available with given (Q, N) is at

present largely insufficient for confirming or disproving this prediction with crystal clear evidence. Yet, we found that a signature of the presence of power-law structures in the empirical data can be observed (at least for some countries) in the conditional distribution $\mathcal{F}_{\text{EMP}}(x|Q > Q_0)$, provided Q_0 is sufficiently large. This yields indirect evidence that $F_{\text{EMP}}(x|Q, N)$ has a power-law tail. It turns out indeed that $\mathcal{F}_{\text{EMP}}(x|Q > Q_0)$ is close to $F_{\text{EMP}}(x|Q_0, \bar{N})$, with $\bar{N} = \mathbb{E}[N|Q > Q_0] \gg 1$, for $x \ll Q_0$ and $Q_0 \gg 1$. Moreover, it is possible to find a range of values for Q_0 where the pool of data contributing to $\mathcal{F}_{\text{EMP}}(x|Q > Q_0)$ is sufficiently large to yield an acceptable signal-to-noise ratio and the domain of $F_{\text{EMP}}(x|Q_0, \bar{N})$ is sufficiently extended to reveal the asymptotic shape of the right tail of the distribution.

In consideration of this, we studied the equations of the model in the large-list limit, *i.e.* the double limit $Q, N \rightarrow \infty$. For pedagogical reasons, we first presented a derivation of the vote distribution in a quenched model, where the original branching-like process is replaced by a supercritical branching process having similar features. The main results of our analysis are that the vote distribution converges quickly to a convolution of single-tree distributions, with weights given by the probabilities of the stopping time of the model, and that this convolution can be discrete or continuous depending on the stopping rule of the model. As a second step we showed that the original branching-like process scales just like a supercritical branching process, hence the solution we found for our quenched model applies *mutatis mutandis* also to the original one. Finally, we presented a derivation of the power-law tail of the vote distribution in the quenched model. The resulting estimate holds, within a reasonable approximation, also for the vote distribution of the original model.

While developing the ideas presented along this exploratory paper, we ran into questions that still look for an answer and represent directions of future research. We list them below in the same order they arise in the text:

- Finland is the only country in the U group of ref. [4] for which $\mathcal{F}_{\text{EMP}}(x|Q > Q_0)$ displays a lognormal behavior independently of Q_0 . It is not clear at present what makes it different from the other countries.
- A relevant difference between eq. (2.1) and eq. (2.4) is that the extinction probability is positive for the former while it vanishes for the latter. In other words, the subspace of finite trees has a positive measure under eq. (2.1). It is not clear whether/how this feature affects the vote distribution.
- We derived the analytic structure of the vote distribution under SR1 in the *minimal* and *maximal ensembles*. It is not clear how to extend the derivation so as to take into account the whole space of trees.
- We computed the distribution of the stopping time by means of numerical simulations. Although we know from general arguments that this is related to a Lévy stable distribution under SR2 and SR3, it is not clear how to go beyond numerical simulations under SR1 in the *minimal* and *maximal ensembles*.
- We calculated the exponential growth rate and proved the martingale property of the branching-like trees introduced by Fortunato and Castellano in a sort of mean-field approximation. Our results hold true up to corrections. It is not clear at present how to improve the estimates.
- A more orthodox derivation of eq. (2.10) than provided in sect. 5 is necessary to ensure a full control of the power-law tail of the vote distribution.
- The representation of the OSTP given in App. B follows from a totally different approach than pursued in sect. 5. It is not clear whether/how it could be used effectively to calculate the coefficient $c(\alpha, n_{\min})$ introduced in eq. (5.22).

To conclude, we observe that a complete understanding of the analytic structure of $F_{\text{FC}}(x|Q, N)$ seems to be a necessary condition to pin down the exact formula relating the microscopic parameters of the word-of-mouth model to the macroscopic parameters of the universal scaling distribution. This will be maybe

a premise to shed light on the ultimate mechanisms lying behind the scaling and universality properties of vote distributions in proportional elections.

Acknowledgments

The computing resources used for our numerical study and the related technical support have been provided by the CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [18]. CRESCO (Computational REsearch centre on COmplex systems) is funded by ENEA and by Italian and European research programmes.

Appendix A Proof of eq. (1.5)

The distribution $\mathcal{F}_{\text{EMP}}(x)$ can be regarded as a convolution of distributions $F_{\text{EMP}}(x|Q, N)$ with weights $p(Q, N)$ and $Q, N = 1, 2, \dots$. We can split the sum over Q into sums over $Q \leq Q_0$ and $Q > Q_0$, namely

$$\begin{aligned}\mathcal{F}_{\text{EMP}}(x) &= \sum_{Q \leq Q_0} \sum_N p(Q, N) F_{\text{EMP}}(x|Q, N) + \sum_{Q > Q_0} \sum_N p(Q, N) F_{\text{EMP}}(x|Q, N) \\ &= \mathcal{F}_{\text{EMP}}^{(Q \leq Q_0)}(x) + \mathcal{F}_{\text{EMP}}^{(Q > Q_0)}(x).\end{aligned}\tag{A.1}$$

Notice, however that neither $\mathcal{F}_{\text{EMP}}^{(Q \leq Q_0)}$ nor $\mathcal{F}_{\text{EMP}}^{(Q > Q_0)}$ are probability densities, in that they fulfill

$$\int_0^\infty dx \mathcal{F}_{\text{EMP}}^{(Q \leq Q_0)}(x) = \sum_{Q \leq Q_0} \sum_N p(Q, N) < 1 \quad \text{and} \quad \int_0^\infty dx \mathcal{F}_{\text{EMP}}^{(Q > Q_0)}(x) = \sum_{Q > Q_0} \sum_N p(Q, N) < 1.\tag{A.2}$$

Since $F_{\text{EMP}}(x|Q, N) \neq 0$ only for $Q/N \leq x \leq Q$, it follows that $\mathcal{F}_{\text{EMP}}^{(Q \leq Q_0)}(x) = 0$ for $x > Q_0$. Hence,

$$\mathcal{F}_{\text{EMP}}(x) = \mathcal{F}_{\text{EMP}}^{(Q > Q_0)}(x), \quad \text{for } x > Q_0.\tag{A.3}$$

Analogously, the distribution $\mathcal{F}_{\text{EMP}}(x|Q > Q_0)$ can be regarded as a convolution of distributions $F_{\text{EMP}}(x|Q, N)$ with weights $p(Q, N|Q > Q_0)$ and $Q > Q_0, N = 1, 2, \dots$, see eqs. (1.6)–(1.7). Since $p(Q, N|Q > Q_0)$ is related to $p(Q, N)$ by

$$p(Q, N) = p(Q, N|Q > Q_0)p(Q > Q_0), \quad \text{for } Q > Q_0,\tag{A.4}$$

it follows that

$$\mathcal{F}_{\text{EMP}}(x) = \mathcal{F}_{\text{EMP}}^{(Q > Q_0)}(x) = p(Q > Q_0) \mathcal{F}_{\text{EMP}}(x|Q > Q_0) < \mathcal{F}_{\text{EMP}}(x|Q > Q_0),\tag{A.5}$$

for $x > Q_0$.

Appendix B OSTP and generalized multiple ζ -values

We can represent $\mathbb{P}\{Z_t = z | Z_{t-1} = \tilde{z}\}$ beyond the asymptotic regime $\tilde{z} \rightarrow \infty$ as an inverse power series in z with coefficient functions amounting to linear combinations of generalized multiple ζ -values. Our derivation assumes $\alpha \in \mathbb{N}$. It is not clear at present whether/how it could be extended to non-integer α .

We start by recalling a simple property of the generating functions of discrete probability distributions. Let $(a_n)_{n=0}^\infty$ and $(b_n)_{n=0}^\infty$ be two sequences of real numbers and let $A(y) = \sum_{n=0}^\infty a_n y^n$ and $B(y) = \sum_{n=0}^\infty b_n y^n$

denote their respective generating functions. Consider also the convolution of $(a_n)_{n=0}^\infty$ and $(b_n)_{n=0}^\infty$, *i.e.* the sequence $(c_n)_{n=0}^\infty$ defined by

$$c_n = \sum_{n_1, n_2=0}^{\infty} \delta_{n, n_1+n_2} a_{n_1} b_{n_2}, \quad (\text{B.1})$$

and let $C(y) = \sum_{n=0}^{\infty} c_n y^n$ denote in turn the generating function of $(c_n)_{n=0}^\infty$. It can be easily checked that

$$C(y) = A(y)B(y). \quad (\text{B.2})$$

Eq. (B.2) can be generalized recursively to the convolution of an arbitrary number of sequences. Since it holds without any restriction on $(a_n)_{n=0}^\infty$ and $(b_n)_{n=0}^\infty$, it applies in particular when $a_n \geq 0$, $b_n \geq 0$ and $A(1) = B(1) = 1$, *i.e.* when $a_n = a(n)$ and $b_n = b(n)$ are discrete probabilities. In this case we have $C(1) = 1$ and thus $c_n = c(n) \geq 0$ is a discrete probability too.

We have already observed that $\mathbb{P}\{Z_t = z | Z_{t-1} = \tilde{z}\}$ is the \tilde{z} -fold convolution of $p_Q(n)$, see eq. (5.7) (first line). Since $G_Q(y) = y_{\min}^n \Phi(y, \alpha, n_{\min}) / \zeta(\alpha, n_{\min})$ is the generating function of $p_Q(n)$, it follows by differentiation that

$$\mathbb{P}\{Z_t = z | Z_{t-1} = \tilde{z}\} = \frac{[\zeta(\alpha, n_{\min})]^{-\tilde{z}}}{z!} \frac{d^z}{dy^z} \Psi(y, \alpha, n_{\min})^{\tilde{z}} \Big|_{y=0}. \quad (\text{B.3})$$

with $\Psi(y, \alpha, n_{\min}) \equiv y^{n_{\min}} \Phi(y, \alpha, n_{\min}) = \sum_{n=n_{\min}}^{\infty} y^n / n^\alpha$. This function is a generalization of the polylogarithm $\text{Li}_\alpha(y)$. Indeed, we have $\text{Li}_\alpha(y) = \Psi(y, \alpha, 1)$. Since the product of polylogarithms can be expressed in terms of multiple polylogarithms, we expect a similar representation to hold also for $\Psi(y, \alpha, n_{\min})^{\tilde{z}}$. In sect. B.1 we recall some basic elements of the theory of multiple polylogarithms and shuffle products (for this we follow ref. [19]). In sect. B.2 we discuss how the theory can be adapted to our case and be used to work out eq. (B.3). We notice incidentally that over the past few years multiple polylogarithms and shuffle products have been gaining popularity in the context of quantum field theory, see ref. [20] for a review. It is funny to see that they can be used to describe complex systems too.

B.1 Review of multiple polylogarithms and shuffle products

The polylogarithm function generalizes the usual logarithm and the Riemann ζ -function. For $\alpha \in \mathbb{N}$ and $y \in \mathbb{C}$ with $|y| \leq 1$ and $(\alpha, y) \neq (1, 1)$, the polylogarithm is defined by the Taylor series

$$\text{Li}_\alpha(y) = \sum_{n=1}^{\infty} \frac{y^n}{n^\alpha}. \quad (\text{B.4})$$

For $\alpha = 1$ and $y \neq 1$ it fulfills $\text{Li}_1(y) = -\log(1-y)$. For $\alpha \geq 2$ and $y = 1$ it fulfills $\text{Li}_\alpha(1) = \zeta(\alpha)$. Polylogarithms can be also defined recursively via differential equations, namely

$$\begin{cases} \frac{d}{dy} \text{Li}_1(y) = \frac{1}{1-y}, \\ \frac{d}{dy} \text{Li}_\alpha(y) = \frac{1}{y} \text{Li}_{\alpha-1}(y), \quad \text{for } \alpha \geq 2, \end{cases} \quad (\text{B.5})$$

with initial condition $\text{Li}_\alpha(0) = 0$, as can be seen by differentiating eq. (B.4) term by term. Recursive integration of eq. (B.5) yields the nested integral representation

$$\begin{aligned} \text{Li}_\alpha(y) &= \int_0^y \frac{dt_1}{t_1} \text{Li}_{\alpha-1} = \int_0^y \frac{dt_1}{t_1} \int_0^{t_1} \frac{dt_2}{t_2} \dots \int_0^{t_{\alpha-2}} \frac{dt_{\alpha-1}}{t_{\alpha-1}} \int_0^{t_{\alpha-1}} \frac{dt_\alpha}{1-t_\alpha} \\ &= \int_{\mathcal{J}_\alpha(y)} \frac{dt_1}{t_1} \frac{dt_2}{t_2} \dots \frac{dt_{\alpha-1}}{t_{\alpha-1}} \frac{dt_\alpha}{1-t_\alpha}, \end{aligned} \quad (\text{B.6})$$

where $\mathcal{J}_\alpha(y)$ is the α -dimensional domain

$$\mathcal{J}_\alpha(y) = \{ (t_1, \dots, t_\alpha) : y > t_1 > t_2 > \dots > t_\alpha > 0 \}. \quad (\text{B.7})$$

The rightmost integral in eq. (B.6) belongs to the class of Chen iterated integrals [21]. Specifically, given the holomorphic 1-forms $\omega_0 = dt/t$ and $\omega_1 = dt/(1-t)$, we can recast eq. (B.6) in the form

$$\text{Li}_\alpha(y) = \int_0^y \underbrace{\omega_0 \circ \dots \circ \omega_0}_{(\alpha-1) \text{ times}} \circ \omega_1 = \int_0^z \omega_0^{\circ(\alpha-1)} \circ \omega_1, \quad (\text{B.8})$$

where the nesting operator \circ is defined recursively by $\int_0^y \phi_1 \circ \dots \circ \phi_k = \int_0^y \phi_1(t) \int_0^t \phi_2 \circ \dots \circ \phi_k$. In the following we let $\tilde{\omega}_\alpha = \omega_0^{\circ(\alpha-1)} \circ \omega_1$ and accordingly we write $\text{Li}_\alpha(y)$ in the more compact notation $\text{Li}_\alpha(y) = \int_0^y \tilde{\omega}_\alpha$.

The product of polylogarithms can be expressed in terms of Chen iterated integrals. As an example, we consider the product $\text{Li}_1(y)\text{Li}_2(y)$ for $y \in \mathbb{R}$ and $0 < y < 1$. This is given by

$$\text{Li}_1(y)\text{Li}_2(y) = \int_{\mathcal{J}_1(y)} \frac{dt_1}{1-t_1} \int_{\mathcal{J}_2(y)} \frac{dt_2}{t_2} \frac{dt_3}{1-t_3} = \int_{\mathcal{J}_1(y) \times \mathcal{J}_2(y)} \frac{dt_1}{1-t_1} \frac{dt_2}{t_2} \frac{dt_3}{1-t_3}. \quad (\text{B.9})$$

The expression on the r.h.s. is not a Chen iterated integral. Nevertheless, the 3-dimensional domain $\mathcal{J}_1(y) \times \mathcal{J}_2(y) = \{ (t_1, t_2, t_3) : y > t_1 > 0 \text{ and } y > t_2 > t_3 > 0 \}$ can be decomposed into nested domains. Indeed, we have

$$\mathcal{J}_1(y) \times \mathcal{J}_2(y) = \mathcal{J}_3(y) \cup \mathcal{J}_3'(y) \cup \mathcal{J}_3''(y), \quad (\text{B.10})$$

up to zero-measure sets, with

$$\mathcal{J}_3'(y) = \{ (t_1, t_2, t_3) : y > t_2 > t_1 > t_3 > 0 \}, \quad (\text{B.11})$$

$$\mathcal{J}_3''(y) = \{ (t_1, t_2, t_3) : y > t_2 > t_3 > t_1 > 0 \}. \quad (\text{B.12})$$

The decomposition of $\mathcal{J}_1(y) \times \mathcal{J}_2(y)$ into the disjoint union of $\mathcal{J}_3(y)$, $\mathcal{J}_3'(y)$ and $\mathcal{J}_3''(y)$ is shown in Fig. 15. Thanks to it, the integral representing $\text{Li}_1(y)\text{Li}_2(y)$ splits into three contributions, each taking the form of a Chen iterated integral, namely

$$\text{Li}_1(y)\text{Li}_2(y) = \int_0^y \omega_1 \circ \omega_0 \circ \omega_1 + 2 \int_0^y \omega_0 \circ \omega_1^2 = \int_0^y (\omega_1 \circ \omega_0 \circ \omega_1 + 2 \omega_0 \circ \omega_1^2). \quad (\text{B.13})$$

The rightmost differential form is obtained by interlacing the 1-form ω_1 of $\text{Li}_1(z)$ with the differential form $\omega_0 \circ \omega_1$ of $\text{Li}_2(z)$ in all possible ways, *i.e.*

$$\omega_1, \omega_0 \circ \omega_1 \longrightarrow \omega_1 \circ \omega_0 \circ \omega_1, \quad \omega_0 \circ \omega_1 \circ \omega_1, \quad \omega_0 \circ \omega_1 \circ \omega_1. \quad (\text{B.14})$$

If we think about ω_0 and ω_1 as playing cards in a deck, the above operation consists in shuffling the sets of cards ω_1 and $\omega_0 \circ \omega_1$ in all possible ways.

The above example can be generalized to the product of an arbitrary number of polylogarithms. Given the 1-forms $\phi_1, \dots, \phi_k, \psi_1, \dots, \psi_n$ with $k, n \geq 1$, we define the shuffle product of $\phi_1 \circ \dots \circ \phi_k$ and $\psi_1 \circ \dots \circ \psi_n$ recursively via

$$\begin{aligned} (\phi_1 \circ \dots \circ \phi_k) \sqcup (\psi_1 \circ \dots \circ \psi_n) &= \phi_1 \circ [(\phi_2 \circ \dots \circ \phi_k) \sqcup (\psi_1 \circ \dots \circ \psi_n)] \\ &\quad + \psi_1 \circ [(\phi_2 \circ \dots \circ \phi_k) \sqcup (\psi_2 \circ \dots \circ \psi_n)], \end{aligned} \quad (\text{B.15})$$

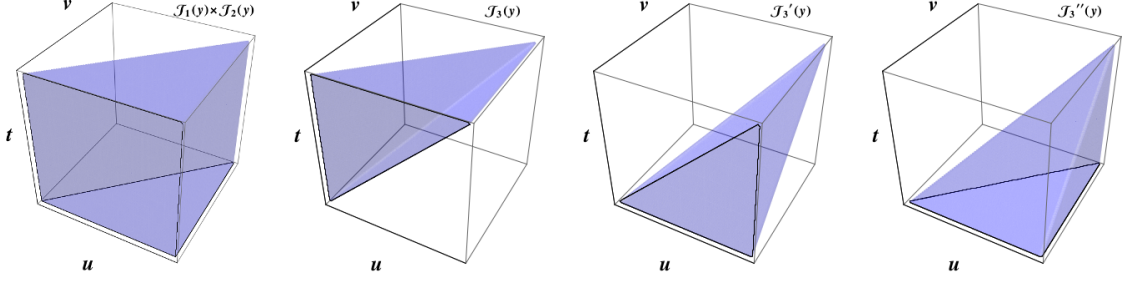


Figure 15 – Decomposition of $\mathcal{J}_1(y) \times \mathcal{J}_2(y)$ into the union of $\mathcal{J}_3(y)$, $\mathcal{J}_3'(y)$ and $\mathcal{J}_3''(y)$.

with $\phi_1 \sqcup \psi_1 = \phi_1 \circ \psi_1 + \psi_1 \circ \phi_1$. For instance, for $k = 1$ and $n = 2$ we have

$$\phi_1 \sqcup (\psi_1 \circ \psi_2) = \phi_1 \circ \psi_1 \circ \psi_2 + \psi_1 \circ \phi_1 \circ \psi_2 + \psi_1 \circ \psi_2 \circ \phi_1, \quad (\text{B.16})$$

for $k = n = 2$ we have

$$\begin{aligned} (\phi_1 \circ \phi_2) \sqcup (\psi_1 \circ \psi_2) &= \phi_1 \circ \phi_2 \circ \psi_1 \circ \psi_2 + \phi_1 \circ \psi_1 \circ \phi_2 \circ \psi_2 \\ &\quad + \phi_1 \circ \psi_1 \circ \psi_2 \circ \phi_2 + \psi_1 \circ \phi_1 \circ \phi_2 \circ \psi_2 \\ &\quad + \psi_1 \circ \phi_1 \circ \psi_2 \circ \phi_2 + \psi_1 \circ \psi_2 \circ \phi_1 \circ \phi_2, \end{aligned} \quad (\text{B.17})$$

and so on and so forth. These examples confirm that the shuffle product interlaces the sets $\{\phi_1, \dots, \phi_k\}$ and $\{\psi_1, \dots, \psi_n\}$ in all possible ways without changing the order of the 1-forms within each set. By arguments similar to those presented above it can be proved in full generality that

$$\int_0^y \phi_1 \circ \dots \circ \phi_k \int_0^y \psi_1 \circ \dots \circ \psi_n = \int_0^y (\phi_1 \circ \dots \circ \phi_k) \sqcup (\psi_1 \circ \dots \circ \psi_n). \quad (\text{B.18})$$

Eq. (B.18) applies immediately to the product of polylogarithms. For instance, we have

$$\text{Li}_\alpha(y) \text{Li}_{\alpha'}(y) = \int_0^y \tilde{\omega}_\alpha \sqcup \tilde{\omega}_{\alpha'} \quad \text{and} \quad [\text{Li}_\alpha(y)]^{\tilde{z}} = \int_0^y \underbrace{\tilde{\omega}_\alpha \sqcup \tilde{\omega}_\alpha \sqcup \dots \sqcup \tilde{\omega}_\alpha}_{\tilde{z} \text{ times}} = \int_0^y \tilde{\omega}_\alpha^{\sqcup \tilde{z}}. \quad (\text{B.19})$$

In particular, each term in the expansion of $[\text{Li}_\alpha(y)]^{\tilde{z}}$ is a Chen iterated integral of the form

$$\text{Li}_{\alpha_1, \dots, \alpha_{\tilde{z}}}(y) = \int_0^y \tilde{\omega}_{\alpha_1} \circ \dots \circ \tilde{\omega}_{\alpha_{\tilde{z}}} = \int_0^y \omega_0^{\circ(\alpha_1-1)} \circ \omega_1 \circ \dots \circ \omega_0^{\circ(\alpha_{\tilde{z}}-1)} \circ \omega_1, \quad (\text{B.20})$$

with the index vector $\{\alpha_1, \dots, \alpha_{\tilde{z}}\}$ fulfilling $\alpha_1 + \dots + \alpha_{\tilde{z}} = \tilde{z}\alpha$ and representing one of many possible ways of interlacing \tilde{z} differential forms $\tilde{\omega}_\alpha$. Reviewing the combinatorics of the indices goes beyond our aims here. We simply define

$$\{\alpha\}_{\tilde{z}} = \underbrace{\{\alpha, \dots, \alpha\}}_{\tilde{z} \text{ times}}, \quad (\text{B.21})$$

and we let $\sqcup\{\alpha\}_{\tilde{z}}$ denote the set of all possible indices $\{\alpha_1, \dots, \alpha_{\tilde{z}}\}$ contributing to the shuffle product. Formally, we have

$$[\text{Li}_\alpha(y)]^{\tilde{z}} = \sum_{\{\alpha_1, \dots, \alpha_{\tilde{z}}\} \in \sqcup\{\alpha\}_{\tilde{z}}} \text{Li}_{\alpha_1, \dots, \alpha_{\tilde{z}}}(y). \quad (\text{B.22})$$

The function $\text{Li}_{\alpha_1, \dots, \alpha_{\tilde{z}}}(y)$ is the multiple polylogarithm of the k th order in one variable. It generalizes the polylogarithm function in the index space. Although eq. (B.20) provides an integral representation of it,

$\text{Li}_{\alpha_1, \dots, \alpha_{\bar{z}}}(y)$ can be equivalently defined as a Taylor series, namely

$$\text{Li}_{\alpha_1, \dots, \alpha_{\bar{z}}}(y) = \sum_{n_1 \geq n_2 \geq \dots \geq n_{\bar{z}} \geq 1} \frac{y^{n_1}}{n_1^{\alpha_1} \dots n_{\bar{z}}^{\alpha_{\bar{z}}}}, \quad \text{for } y \in \mathbb{C} \text{ and } |y| \leq 1, \quad (\text{B.23})$$

with $\alpha_1, \dots, \alpha_{\bar{z}} \geq 1$ and $(y, \alpha_1) \neq (1, 1)$. By differentiating this series term by term, we get an alternative representation of $\text{Li}_{\alpha_1, \dots, \alpha_{\bar{z}}}(y)$ as the solution of the recursive differential equations

$$\begin{cases} \frac{d}{dy} \text{Li}_{1, \alpha_2, \dots, \alpha_{\bar{z}}}(y) = \frac{1}{1-y} \text{Li}_{\alpha_2, \dots, \alpha_{\bar{z}}}, \\ \frac{d}{dy} \text{Li}_{\alpha_1, \alpha_2, \dots, \alpha_{\bar{z}}}(y) = \frac{1}{y} \text{Li}_{\alpha_1-1, \alpha_2, \dots, \alpha_{\bar{z}}}(y), \quad \text{for } \alpha_1 \geq 2, \end{cases} \quad (\text{B.24})$$

with initial condition $\text{Li}_{\alpha_1, \dots, \alpha_{\bar{z}}}(0) = 0$. Eq. (B.20) is then obtained by recursively integrating eq. (B.24).

We conclude this short review by introducing the multiple ζ -function

$$\zeta(\alpha_1, \dots, \alpha_{\bar{z}}) = \sum_{n_1 \geq n_2 \geq \dots \geq n_{\bar{z}} \geq 1} \frac{1}{n_1^{\alpha_1} \dots n_{\bar{z}}^{\alpha_{\bar{z}}}}, \quad \text{for } \alpha_1 \geq 2 \text{ and } \alpha_2, \dots, \alpha_{\bar{z}} \geq 1. \quad (\text{B.25})$$

This function represents a multi-dimensional generalization of the ordinary Riemann ζ -function and fulfills $\zeta(\alpha_1, \dots, \alpha_{\bar{z}}) = \text{Li}_{\alpha_1, \dots, \alpha_{\bar{z}}}(1)$ for $\alpha_1 \geq 2$.

B.2 Adaptation to the OSTP

The function $\Psi(y, \alpha, n_{\min})$ can be regarded as an upper incomplete (or lower-truncated) polylogarithm, *i.e.* it satisfies the relation

$$\Psi(y, \alpha, n_{\min}) = \text{Li}_{\alpha}(y) - \sum_{j=0}^{n_{\min}-1} \frac{y^j}{j^{\alpha}}. \quad (\text{B.26})$$

From eq. (B.5) and recalling that $\sum_{k=0}^s y^k = (1 - y^{s+1})/(1 - y)$ for $0 < y < 1$, we see that $\Psi(y, \alpha, n)$ fulfills the recursive differential equations

$$\begin{cases} \frac{d}{dy} \Psi(y, 1, n_{\min}) = \frac{y^{n_{\min}-1}}{1-y}, \\ \frac{d}{dy} \Psi(y, \alpha, n_{\min}) = \frac{1}{y} \Psi(y, \alpha-1, n_{\min}), \quad \text{for } \alpha \geq 2, \end{cases} \quad (\text{B.27})$$

with initial condition $\Psi(0, \alpha, n_{\min}) = 0$. As a consequence, it can be represented as a Chen iterated integral,

$$\Psi(y, \alpha, n_{\min}) = \int_0^y \underbrace{\omega_0 \circ \dots \circ \omega_0}_{\alpha-1 \text{ times}} \circ \omega_{1, n_{\min}} = \int_0^y \omega_0^{\circ(\alpha-1)} \circ \omega_{1, n_{\min}} = \int_0^y \tilde{\omega}_{\alpha, n_{\min}}, \quad (\text{B.28})$$

where we let $\omega_{1, n_{\min}} = t^{n_{\min}-1} dt / (1-t)$ and $\tilde{\omega}_{\alpha, n_{\min}} = \omega_0^{\alpha-1} \circ \omega_{1, n_{\min}}$. Since the only difference between eq. (B.28) and eq. (B.8) is that ω_1 is replaced by $\omega_{1, n_{\min}}$, we conclude that $\Psi(y, \alpha, n_{\min})^{\bar{z}}$ can be analogously represented as the integral of the shuffle product of the corresponding nested differential forms, *i.e.*

$$[\Psi(y, \alpha, n_{\min})]^{\bar{z}} = \int_0^y \underbrace{\tilde{\omega}_{\alpha, n_{\min}} \sqcup \tilde{\omega}_{\alpha, n_{\min}} \sqcup \dots \sqcup \tilde{\omega}_{\alpha, n_{\min}}}_{\bar{z} \text{ times}} = \int_0^y \tilde{\omega}_{\alpha, n_{\min}}^{\sqcup \bar{z}}. \quad (\text{B.29})$$

As previously, each term in the expansion of $\omega_{\alpha, n_{\min}}^{\sqcup \tilde{z}}$ gives rise to a Chen iterated integral

$$\begin{aligned}\Psi(y, \{\alpha_1, \dots, \alpha_{\tilde{z}}\}, n_{\min}) &= \int_0^y \tilde{\omega}_{\alpha_1, n_{\min}} \circ \dots \circ \tilde{\omega}_{\alpha_{\tilde{z}}, n_{\min}} \\ &= \int_0^y \omega_0^{\circ(\alpha_1-1)} \circ \omega_{1, n_{\min}} \circ \dots \circ \omega_0^{\circ(\alpha_{\tilde{z}}-1)} \circ \omega_{1, n_{\min}},\end{aligned}\quad (\text{B.30})$$

with the index vector $\{\alpha_1, \dots, \alpha_{\tilde{z}}\}$ fulfilling $\alpha_1 + \dots + \alpha_{\tilde{z}} = \tilde{z}\alpha$ and representing one of many possible ways of interlacing \tilde{z} differential forms $\tilde{\omega}_{\alpha, n_{\min}}$. Formally, we have

$$[\Psi(y, \alpha, n_{\min})]^{\tilde{z}} = \sum_{\{\alpha_1, \dots, \alpha_{\tilde{z}}\} \in \sqcup \{\alpha\}_{\tilde{z}}} \Psi(y, \{\alpha_1, \dots, \alpha_{\tilde{z}}\}, n_{\min}). \quad (\text{B.31})$$

The function $\Psi(y, \{\alpha_1, \dots, \alpha_{\tilde{z}}\}, n_{\min})$ is a lower-truncated multiple polylogarithm. It generalizes the function $\Psi(y, \alpha, n_{\min})$ in the index space. Upon differentiating eq. (B.30), we see that $\Psi(y, \{\alpha_1, \dots, \alpha_{\tilde{z}}\}, n_{\min})$ fulfills the recursive differential equations

$$\begin{cases} \frac{d}{dy} \Psi(y, \{1, \alpha_2, \dots, \alpha_{\tilde{z}}\}, n_{\min}) = \frac{y^{n_{\min}-1}}{1-y} \Psi(y, \{\alpha_2, \dots, \alpha_{\tilde{z}}\}, n_{\min}), \\ \frac{d}{dy} \Psi(y, \{\alpha_1, \dots, \alpha_{\tilde{z}}\}, n_{\min}) = \frac{1}{y} \Psi(y, \{\alpha_1 - 1, \alpha_2, \dots, \alpha_{\tilde{z}}\}, n_{\min}), \quad \text{for } \alpha_1 \geq 2, \end{cases} \quad (\text{B.32})$$

with initial condition $\Psi(0, \{\alpha_1, \dots, \alpha_{\tilde{z}}\}, n_{\min}) = 0$. Similar to the multiple polylogarithm, also the function $\Psi(y, \{\alpha_1, \dots, \alpha_k\}, n_{\min})$ can be represented as a power series, *i.e.*

$$\Psi(y, \{\alpha_1, \dots, \alpha_{\tilde{z}}\}, n_{\min}) = \sum_{n_1=\tilde{z}n_{\min}}^{\infty} \sum_{n_2=(\tilde{z}-1)n_{\min}}^{n_1-n_{\min}} \dots \sum_{n_{\tilde{z}}=n_{\min}}^{n_{\tilde{z}-1}-n_{\min}} \frac{y^{n_1}}{n_1^{\alpha_1} \dots n_{\tilde{z}}^{\alpha_{\tilde{z}}}}. \quad (\text{B.33})$$

The reader can easily check that the r.h.s of eq. (B.33) fulfills eq. (B.32) for $\alpha_1 \geq 2$. We want to prove that it fulfills eq. (B.32) also for $\alpha_1 = 1$. To this aim, we let $\tilde{\Psi}(y, \{\alpha_1, \dots, \alpha_{\tilde{z}}\}, n_{\min})$ denote the r.h.s. of eq. (B.33). Moreover, to simplify the notation, we let

$$\tilde{\Psi}(y, \{1, \alpha_2, \dots, \alpha_{\tilde{z}}\}, n_{\min}) = \sum_{n_1=\tilde{z}n_{\min}}^{\infty} c_{n_1} \frac{y^{n_1}}{n_1}, \quad (\text{B.34})$$

with the coefficient c_{n_1} being given by

$$c_{n_1} = \sum_{n_2=(\tilde{z}-1)n_{\min}}^{n_1-n_{\min}} \dots \sum_{n_{\tilde{z}}=n_{\min}}^{n_{\tilde{z}-1}-n_{\min}} \frac{1}{n_2^{\alpha_2} \dots n_{\tilde{z}}^{\alpha_{\tilde{z}}}}. \quad (\text{B.35})$$

Now, we have

$$\begin{aligned}(1-y) \frac{d}{dy} \tilde{\Psi}(y, \{1, \alpha_2, \dots, \alpha_{\tilde{z}}\}, n_{\min}) &= \sum_{n_1=\tilde{z}n_{\min}}^{\infty} c_{n_1} y^{n_1-1} - \sum_{n_1=\tilde{z}n_{\min}}^{\infty} c_{n_1} y^{n_1} \\ &= c_{\tilde{z}n_{\min}} y^{\tilde{z}n_{\min}-1} + \sum_{n_1=\tilde{z}n_{\min}}^{\infty} (c_{n_1+1} - c_{n_1}) y^{n_1}.\end{aligned}\quad (\text{B.36})$$

The coefficient $c_{n_1+1} - c_{n_1}$ is obtained from eq. (B.35) by removing the sum over n_2 and by calculating the rest at $n_2 = n_1 + 1 - n_{\min}$, *i.e.*

$$c_{n_1+1} - c_{n_1} = \sum_{n_3=(\tilde{z}-2)n_{\min}}^{n_1+1-2n_{\min}} \sum_{n_4=(\tilde{z}-3)n_{\min}}^{n_3-n_{\min}} \dots \sum_{n_{\tilde{z}}=n_{\min}}^{n_{\tilde{z}-1}-n_{\min}} \frac{1}{(n_1+1-n_{\min})^{\alpha_2} n_3^{\alpha_3} \dots n_{\tilde{z}}^{\alpha_{\tilde{z}}}}. \quad (\text{B.37})$$

Inserting this expression into the previous equation yields

$$\begin{aligned}
(1-y) \frac{d}{dy} \tilde{\Psi}(y, \{\alpha_1, \dots, \alpha_{\tilde{z}}\}, n_{\min}) &= c z n_{\min} y^{\tilde{z} n_{\min} - 1} + \sum_{n_1 = \tilde{z} n_{\min}}^{\infty} \sum_{n_3 = (\tilde{z}-2)n_{\min}}^{n_1 + 1 - 2n_{\min}} \cdots \sum_{n_{\tilde{z}} = n_{\min}}^{n_{\tilde{z}-1} - n_{\min}} \frac{y^{n_1}}{(n_1 + 1 - n_{\min})^{\alpha_2} n_3^{\alpha_3} \dots n_{\tilde{z}}^{\alpha_{\tilde{z}}}} \\
&= y^{n_{\min} - 1} \left\{ \sum_{n_2 = (\tilde{z}-1)n_{\min}}^{\infty} \sum_{n_3 = (\tilde{z}-2)n_{\min}}^{n_2 - n_{\min}} \cdots \sum_{n_{\tilde{z}} = n_{\min}}^{n_{\tilde{z}-1} - n_{\min}} \frac{y^{n_2}}{n_2^{\alpha_2} n_3^{\alpha_3} \dots n_{\tilde{z}}^{\alpha_{\tilde{z}}}} \right\} \\
&= y^{n_{\min} - 1} \tilde{\Psi}(y, \{\alpha_2, \dots, \alpha_{\tilde{z}}\}, n_{\min}). \tag{B.38}
\end{aligned}$$

It follows that $\tilde{\Psi}(y, \{1, \alpha_2, \dots, \alpha_{\tilde{z}}\}, n_{\min}) = \Psi(y, \{1, \alpha_2, \dots, \alpha_{\tilde{z}}\}, n_{\min})$.

Just as the multiple ζ -function $\zeta(\alpha_1, \dots, \alpha_{\tilde{z}})$ is obtained from the multiple polylogarithm $\text{Li}_{\alpha_1, \dots, \alpha_{\tilde{z}}}(y)$ by taking the latter at $y = 1$, we define a truncated multiple ζ -function $\zeta(\{\alpha_1, \dots, \alpha_{\tilde{z}}\}, n_{\min})$ from the truncated multiple polylogarithm $\Psi(y, \{\alpha_1, \dots, \alpha_{\tilde{z}}\}, n_{\min})$ by taking the latter at $y = 1$, namely

$$\zeta(\{\alpha_1, \dots, \alpha_{\tilde{z}}\}, n_{\min}) = \sum_{n_1 = \tilde{z} n_{\min}}^{\infty} \sum_{n_2 = (\tilde{z}-1)n_{\min}}^{n_1 - n_{\min}} \cdots \sum_{n_{\tilde{z}} = n_{\min}}^{n_{\tilde{z}-1} - n_{\min}} \frac{1}{n_1^{\alpha_1} \dots n_{\tilde{z}}^{\alpha_{\tilde{z}}}}. \tag{B.39}$$

This function is a multi-dimensional generalization of the Hurwitz ζ -function $\zeta(\alpha, n_{\min})$.

By means of the above formalism we can work out eq. (B.3). Since $y^{\tilde{z} n_{\min}}$ is the monomial of lowest degree contributing to $\Psi(y, \{\alpha_1, \dots, \alpha_{\tilde{z}}\}, n_{\min})$, we have

$$\begin{aligned}
\frac{1}{z!} \frac{d^z}{dy^z} \Psi(y, \{\alpha_1, \dots, \alpha_{\tilde{z}}\}, n_{\min}) \Big|_{y=0} &= \frac{\theta_{z, \tilde{z} n_{\min}}}{z!} \frac{d^z}{dy^z} \left\{ \frac{y^z}{z^{\alpha_1}} \sum_{n_2 = (\tilde{z}-1)n_{\min}}^{z - n_{\min}} \cdots \sum_{n_{\tilde{z}} = n_{\min}}^{n_{\tilde{z}-1} - n_{\min}} \frac{1}{n_2^{\alpha_2} \dots n_{\tilde{z}}^{\alpha_{\tilde{z}}}} \right\} \\
&= \theta_{z, \tilde{z} n_{\min}} \frac{\zeta(z, \{\alpha_2, \dots, \alpha_{\tilde{z}}\}, n_{\min})}{z^{\alpha_1}}, \tag{B.40}
\end{aligned}$$

with

$$\zeta(z, \{\alpha_2, \dots, \alpha_{\tilde{z}}\}, n_{\min}) = \sum_{n_2 = (\tilde{z}-1)n_{\min}}^{z - n_{\min}} \cdots \sum_{n_{\tilde{z}} = n_{\min}}^{n_{\tilde{z}-1} - n_{\min}} \frac{1}{n_2^{\alpha_2} \dots n_{\tilde{z}}^{\alpha_{\tilde{z}}}}, \tag{B.41}$$

representing a doubly truncated multiple ζ -function. It follows that

$$\mathbb{P}\{Z_t = z \mid Z_{t-1} = \tilde{z}\} = \theta_{z, \tilde{z} n_{\min}} \sum_{\{\alpha_1, \dots, \alpha_{\tilde{z}}\} \in \mathbb{W}\{\alpha\}_{\tilde{z}}} \frac{1}{z^{\alpha_1}} \frac{\zeta(z, \{\alpha_2, \dots, \alpha_{\tilde{z}}\}, n_{\min})}{\zeta(\alpha, n_{\min})^{\tilde{z}}}. \tag{B.42}$$

The presence of $\theta_{z, \tilde{z} n_{\min}}$ is perfectly natural, since each of the \tilde{z} vertices on the $(t-1)$ th level of a quenched tree generates at least n_{\min} vertices. It can be easily shown that $\alpha_1 \geq \alpha$ for each $\{\alpha_1, \dots, \alpha_{\tilde{z}}\} \in \mathbb{W}\{\alpha\}_{\tilde{z}}$. Suppose indeed that you have a deck of playing cards with only two different types of cards, namely ω_0 and ω_1 . The deck is made of $(\alpha-1)\tilde{z}$ cards of type ω_0 and \tilde{z} cards of type ω_1 . These are initially stacked in iterated groups $\underbrace{\omega_0, \dots, \omega_0}_{\alpha-1 \text{ times}}, \omega_1$, namely

$$\underbrace{\omega_0, \dots, \omega_0}_{\alpha-1 \text{ times}}, \omega_1, \underbrace{\omega_0, \dots, \omega_0}_{\alpha-1 \text{ times}}, \omega_1, \dots, \underbrace{\omega_0, \dots, \omega_0}_{\alpha-1 \text{ times}}, \omega_1. \tag{B.43}$$

The basic rule of the shuffle product states that upon shuffling the deck, each card of type ω_1 must always stay to the right of all cards ω_0 belonging to its group. It follows that each allowed reshuffling of cards has at least $\alpha - 1$ cards of type ω_0 on the left, *i.e.* $\alpha_1 - 1 \geq \alpha - 1$. Hence, eq. (B.42) represents the OSTP as a series of inverse powers $z^{-\alpha_1}$ with degree $\alpha_1 \geq \alpha$ and non-negative modulating coefficient functions $\zeta(z, \{\alpha_2, \dots, \alpha_{\tilde{z}}\}, n_{\min}) / \zeta(\alpha, n_{\min})^{\tilde{z}}$. A comparison between eq. (B.42) and eq. (5.17) inspires the conjecture

$$\lim_{\tilde{z} \rightarrow \infty} \lim_{z \rightarrow \infty} \frac{1}{\tilde{z}} \sum_{\substack{\{\alpha_1, \dots, \alpha_{\tilde{z}}\} \in \mathbb{W}\{\alpha\}_{\tilde{z}} \\ \text{with } \alpha_1 = \alpha}} \frac{\zeta(z, \{\alpha_2, \dots, \alpha_{\tilde{z}}\}, n_{\min})}{\zeta(\alpha, n_{\min})^{\tilde{z}-1}} = 1. \quad (\text{B.44})$$

It is not clear whether this appealing formula is correct at all, since eq. (5.17) has been derived under the hypothesis $\alpha \neq 1, 2, 3, \dots$ while eq. (B.42) has been obtained for integer α . We just notice that each term on the l.h.s. has exactly $\tilde{z} - 1$ nested sums at numerator and $\tilde{z} - 1$ independent sums at denominator. We do not even know how eq. (B.42) could be extended to non-integer values of α .

References

- [1] S. Fortunato and C. Castellano. Scaling and universality in proportional elections. *Phys. Rev. Lett.*, 99(13):138701, 2007.
- [2] K. Burghardt, W. Rand, and M. Girvan. Competing opinions and stubbornness: connecting models to data. *ArXiv e-prints*, 2014, 1411.7415.
- [3] S. Fortunato and C. Castellano. Word of mouth and universal voting behavior in proportional elections. Available at <http://sites.google.com/site/santofortunato/Oxford.ppt?attredirects=0>. (accessed: 01-11-2015). Invited seminar for the Complex Adaptive Systems Group (CASG), Hilary Term 2007, Saïd Business School, University of Oxford, UK (Feb. 2007).
- [4] A. Chatterjee, M. Mitrović, and S. Fortunato. Universality in voting behavior: an empirical analysis. *Scientific Reports*, 3, 2013, 1212.2142.
- [5] B. Mandelbrot. Information Theory and Psycholinguistics. In B. B. Wolman and E. Nagel, *Scientific Psychology*, Basic Books, 1965.
- [6] R. Otter. The multiplicative process. *Ann. Math. Statist.*, 20(2):206–224, 06 1949.
- [7] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 2. Wiley, January 1968.
- [8] T. E. Harris. *The theory of branching processes*. Die Grundlehren der Mathematischen Wissenschaften, Bd. 119. Springer-Verlag, 1963.
- [9] K. B. Athreya and P. E. Ney. *Branching Processes*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, reprint of the original 1972 edition, 2012.
- [10] D.S. Lee, J.S. Kim, B. Kahng, and D. Kim. Scale-free random branching tree in supercritical phase. *J. Phys. A:Math. Theor.*, 40:7139–7149, 2007.
- [11] A. V. Manohar. An Introduction to spin dependent deep inelastic scattering. In *Lake Louise Winter Institute: Symmetry and Spin in the Standard Model Lake Louise, Alberta, Canada, February 23-29, 1992*, 1992, hep-ph/9204208.

- [12] G. Samorodnitsky and M. S. Taquq. *Stable non-Gaussian random processes : stochastic models with infinite variance*. Chapman & Hall, New York, 1994. Reprint by CRC Press (2000).
- [13] V. V. Uchaikin and V. M. Zolotarev. *Chance and stability: stable distributions and their applications*. VSP International Science Publishers, Utrecht, 1999.
- [14] B. Roehner and P. Winiwarter. Aggregation of independent Paretian random variables. *Adv. Appl. Prob.*, 17:465–469, 1985.
- [15] A. Erdélyi et al. *Higher transcendental functions*, volume 1. McGraw Hill, New York, 1953.
- [16] B. Gnedenko and V. Kolmogorov. *Limit distributions for sums of Independent Random Variables*. Addison–Wesley, Reading, Mass., 1954.
- [17] J. P. Nolan. Stable distributions. Available at <http://academic2.american.edu/~jpnolan/stable/chap1.ps>, 2009. (accessed: 01-11-2015).
- [18] G. Ponti et al. The role of medium size facilities in the HPC ecosystem: the case of the new CRESCO4 cluster integrated in the ENEAGRID infrastructure. In *Proceedings of the 2014 International Conference on High Performance Computing and Simulation - HPCS2014*, number 6903807, pages 1030–1033, 2014.
- [19] M. Waldschmidt. Multiple polylogarithms: an introduction. In *Number Theory and Discrete Mathematics (Trends in Mathematics)*. A. K. Agarwal et al. editors. Birkhäuser; Softcover reprint of the original 1st ed. 2002 edition, 2012.
- [20] F. Brown. Iterated integrals in quantum field theory. In *Geometric and Topological Methods for Quantum Field. Theory Proceedings of the 2009 Villa de Leyva Summer School*, pages 188–240, 2013.
- [21] K. T. Chen. Iterated path integrals. *Bull. Amer. Math. Soc.*, 83(5):831–879, 1977.